

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Errorless Learning of Prospective Memory Tasks: An Experimental Investigation in People with Memory Disorders.

Fish, Jess

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

This electronic theses or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Title:Errorless Learning of Prospective Memory Tasks: An Experimental Investigation in People with Memory Disorders.

Author:Jess Fish

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENSE AGREEMENT



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. <http://creativecommons.org/licenses/by-nc-nd/3.0/>

You are free to:

- Share: to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

**Volume I: Main Research Project and
Service-Related Research Project**

Jessica E. Fish, Ph.D.

Thesis submitted for the Doctorate in
Clinical Psychology, King's College
London Institute of Psychiatry.

May 2013

Acknowledgements

Many thanks are due to Professor Robin Morris and Professor Mike Kopelman for their supervision of this project. I am extremely lucky to have been in receipt of such thoroughly expert guidance! I would also like to express my gratitude to Professor Barbara A. Wilson, whose work provided the impetus for this project, and who is an ongoing source of inspiration and sage advice. The sterling help of Dr Eli Jaldow and Dr Kate Humphreys of the Neuropsychiatry and Memory Disorders Service, particularly in relation to recruitment, is also much appreciated. Thanks also to Drs Patrick Smith and Lidia Yáguez for reviewing the proposal, the Psychological Medicine CAG for supporting the project, Jenny Liebscher and her team for enabling navigation of REC and R&D approval procedures, Dr Andrew Brand for programming the experimental task, Dr Daniel Stahl for statistical advice, and Dr Tom Manly for his ongoing collaboration and for many interesting discussions, at least some of which were about Prospective Memory. I also very much appreciate the essential contribution made by the participants who kindly volunteered their time towards this project.

I am grateful to Dr Anouk Houdijk for her supervision of the Service-Related Research Project, and to Professor Stephen Scott, Dr Carmen Pinto, and Dr Matt Woolgar for their feedback, and Toni Cole for administrative support.

I would also like to acknowledge all seven of the Clinical Psychologists that I have been fortunate to be supervised by during my training: Rachael Buxey, Dr Sarah Crawford, Dr Antonia Dittner, Dr Anouk Houdijk, Professor Myra Hunter, Dr Eli Jaldow, and Professor Robin Morris. They have made highly significant contributions to this thesis either directly through supervising projects and case studies, indirectly through helping to develop my clinical and general thinking skills, or both (Fish's exact test, $p < .001$).

Dr Benjamin Boecking merits many thanks for his thoughtful comments on this thesis, and for his careful proofreading. The excellent movie recommendations (or 'Filmempfehlungen', if you will) are also much appreciated. Thanks also to the rest of my year group, and especially Claire, Jon, and Harriet, for helping me to stay more or less sane through the process of clinical training; and to the core team and admin staff particularly Sue Rutter, Carole Barnham and Mark Ballham for making sure that all things DClinPsy run so smoothly.

Finally, I acknowledge the National Health Service for funding this work, through its financial investment in Clinical Psychology training.

Summary of Contents

Errorless Learning of Prospective Memory Tasks: An Experimental Investigation in People with Memory Disorders.	4
An Audit of Referrals Accepted by a Tier 4 CAMHS Team Pre- and Post-Publication of NICE Guidance for Conduct Problems.....	94

Main Research Project

Errorless Learning of Prospective Memory Tasks: An Experimental Investigation in People with Memory Disorders.

Supervisors

Professor Robin Morris

King's College London Institute of Psychiatry, and King's
College Hospital NHS Foundation Trust

Professor Mike Kopelman

King's College London Institute of Psychiatry, and South
London and Maudsley NHS Foundation Trust

Discussant

Dr Tom Manly

Medical Research Council Cognition and Brain Sciences
Unit, Cambridge

Table of Contents

List of Tables	9
List of Figures	10
List of Abbreviations	11
1. Abstract	12
2. Introduction	13
2.1. A Working Definition of Prospective Memory	13
2.2. Systems for Categorising PM tasks	14
2.3. The Measurement of Prospective Memory	15
2.3.1. Using Standardised Measures of PM	15
2.3.2. Assessing Component Processes of Prospective Remembering.....	15
2.3.3. Measuring Performance on “Real” PM Tasks	16
2.3.4. Using Self and/or Informant Ratings.....	16
2.3.5. Using Laboratory-Based Tasks Derived from the Literature.....	16
2.3.6. The Ecological Validity of PM Assessments	18
2.3.7. Recommendations for Measuring PM	19
2.4. Theoretical Perspectives on Prospective Memory	20
2.4.1. Hypothesised Stages of PM.....	21
2.4.2. Cognitive Requirements of PM Task Completion	21
2.4.3. Models Specifically Examining Retrieval Processes in PM.....	22
2.4.3.1. The Multiprocess and Preparatory Attention and Memory Models	22
2.4.3.2. Factors that Modulate PM Retrieval.....	24
2.4.4. Metacognitive Aspects of Prospective Memory	25
2.4.5. Neuroscience Perspectives on PM.....	26
2.5. Prospective Memory in Neurological Disorders	28
2.6. Rehabilitation of Prospective Memory	29
2.6.1. Restorative Approaches	30

2.6.2.	Compensatory Approaches.....	30
2.6.2.1.	Task-Level Strategies for PM.....	31
2.6.2.2.	Process-Level Strategies Targeting Executive Demands.....	31
2.6.2.3.	Process-Level Strategies Targeting Mnemonic Demands.....	32
2.7.	Errorless Learning in Memory Rehabilitation	33
2.7.1.	Origins of Errorless Learning	33
2.7.2.	Techniques for Errorless Learning.....	34
2.7.3.	Applications of Errorless Learning Methods.....	35
2.7.4.	Mechanisms of the Errorless Learning Advantage	38
2.7.5.	Applying Errorless Learning to Prospective Memory tasks	41
2.8.	Aims of the Study.....	42
3.	Method	44
3.1.	Ethical Review	44
3.2.	Participants	44
3.2.1.	Recruitment	44
3.2.2.	Eligibility Criteria	44
3.2.3.	Sample Characteristics.....	45
3.2.3.1.	Demographics	45
3.2.3.2.	Aetiological Information	45
3.2.3.3.	Definition of Memory Impairment	45
3.3.	Study Design and Counterbalancing	47
3.4.	Procedure.....	48
3.4.1.	Overview	48
3.4.2.	Encoding Procedures	49
3.4.2.1.	Errorless Learning	50
3.4.2.2.	Errorful Learning	50
3.5.	Measures.....	51
3.5.1.	Prospective Memory Paradigm.....	51
3.5.1.1.	Ongoing Task.....	51

3.5.1.2. PM tasks.....	52
3.5.2. Background Neuropsychological Measures.....	54
3.5.3. Questionnaires.....	54
3.6. Power Calculation	55
3.7. Planned Statistical Analyses.....	55
4. Results.....	57
4.1. Preliminary Analyses Concerning the Efficacy of the Experimental Manipulations.....	57
4.1.1. Determining an Appropriate Analysis Strategy.....	57
4.1.2. The Occurrence of Errors across Encoding Conditions and PM Tasks	57
4.1.3. Exposure to Correct Information across Encoding Conditions and PM Tasks	58
4.1.4. Retrospective Memory for PM Task Content	59
4.2. Primary Analysis: The Impact of Encoding Condition on Prospective Memory Accuracy.....	59
4.2.1. PM Accuracy as a Function of Encoding and Task Type.....	59
4.2.2. Time-based PM Errors by Encoding Condition	61
4.3. Ongoing Task Performance across the Experimental Conditions.....	61
4.3.1. Determining an Appropriate Analysis Strategy.....	61
4.3.2. Repeated Measures MANOVA on OT performance	61
4.4. PM Awareness Variables across Experimental Conditions	62
4.4.1. Determining an Appropriate Analysis Strategy.....	62
4.4.2. Repeated Measures MANOVA on Awareness variables.....	62
5. Discussion.....	65
5.1. Summary of Findings.....	65
5.2. Relation of the Findings to the Literature	66
5.2.1. The Results in Relation to Models of Prospective Memory.....	66
5.2.2. The Results in Relation to Research on Errorless Learning.....	67
5.2.3. The Results in Relation to Research on Metacognition	69
5.3. Limitations of the Study, and Suggested Modifications in Future Research	70
5.3.1. Sample Size and Associated Shortcomings	70

5.3.2.	Limitations Related to Measurement Decisions	71
5.4.	Future Directions	72
5.4.1.	Studying Errorless Learning of Naturalistic Prospective Memory Tasks.....	73
5.4.2.	Studying Naturalistic Encoding Errors.....	73
5.4.3.	Disseminating the Findings and Practical Applications of the Results.....	74
6.	References	76
7.	Appendices.....	91
7.1.	Procedures for Determining Mental Capacity to Consent.....	91
7.2.	Background Information on Study Participants.....	93

List of Tables

Table 1.1. Summary of the advantages and disadvantages of different approaches to measuring Prospective Memory.....	17
Table 1.2. Task administration order. Note that grey/white shading refers to learning condition, whereby grey shading indicates Errorless Learning and white control.....	48
Table 1.3. Rates of errors and prompts by encoding condition and task type (M [SD]).	58
Table 1.4. Descriptive statistics for PM performance and the Errorless Learning Advantage. ..	60
Table 1.5. Descriptive statistics (M, [SEM], 95% Confidence Interval) for Ongoing Task variables according to experimental condition.....	62
Table 1.6. Descriptive Statistics for Awareness variables by Encoding condition	63
Table 1.7. Table showing demographic, medical, neuropsychological and questionnaire data for each participant, and group summaries where relevant.	93

List of Figures

Figure 1.1. Schematic diagram showing stages involved in completion of Prospective Memory tasks and associated cognitive operations.	22
Figure 1.2. Flow chart detailing procedures for Errorless and Errorful Encoding	51
Figure 1.3. Illustration of the trial sequence and response format of the PM task.....	53
Figure 1.4. Mean EBPM and TBPM accuracy in Errorless and Errorful conditions.....	60
Figure 1.5. Mean post-test awareness of PM Targets plotted by Encoding Condition and PM Task Type.	64

List of Abbreviations

ABI	Acquired Brain Injury
AD	Alzheimer's Disease
ADHD	Attention Deficit Hyperactivity Disorder
ANOVA	Analysis of Variance
BMIPB	BIRT Memory and Information Processing Battery
BOLD	Blood Oxygen Level Dependent
CAMPROMPT	Cambridge Prospective Memory Test
CAPM	Comprehensive Assessment of Prospective Memory
EBIQ	European Brain Injury Questionnaire
EBPM	Event-based Prospective Memory
EL	Errorless Learning
ELA	Errorless Learning Advantage
ERP	Event-related Potential
fMRI	functional Magnetic Resonance Imaging
FSIQ	Full-Scale Intelligence Quotient
HADS	Hospital Anxiety and Depression Scale
MRI	Magnetic Resonance Imaging
MTL	Medial Temporal Lobe
ONS	Office for National Statistics
PFC	Prefrontal Cortex
PM	Prospective Memory
PRMQ	Prospective and Retrospective Memory Questionnaire
RBMT	Rivermead Behavioural Memory Test
SCOLP	Speed and Capacity of Language Processing Test
SR	Spaced Retrieval
TBI	Traumatic Brain Injury
TBPM	Time-based Prospective Memory
tDCS	Transcranial Direct Current Stimulation
TOPF	Test of Premorbid Functioning
VC	Vanishing Cues
WM	Working Memory
WASI/WASI-II	Wechsler Abbreviated Scale of Intelligence (and second version)
WTAR	Wechsler Test of Adult Reading

1. Abstract

Prospective Memory (PM), or the ability to act upon delayed intentions, is cognitively complex as it requires a combination of mnemonic, attentional and executive abilities. PM tasks can be particularly challenging for people with cognitive impairment, and it is important to identify effective means of rehabilitation. Errorless Learning (EL) is an encoding technique that results in superior recall and recognition memory performance compared with 'errorful' learning in people with memory impairment. This so-called 'Errorless Learning advantage' (ELA) has been attributed to implicit memory processes (Page et al., 2006), and there is a basis for predicting a similar beneficial effect on PM performance. However, PM tasks vary in their retrieval demands, some involving environmentally-cued retrieval of a cue-action association (referred to as Event-based PM tasks), and some requiring self-cued retrieval of the action to be performed (referred to as Time-based PM tasks). Event-based PM performance may, therefore, be seen to rely more upon mnemonic processes, and Time-based PM performance on more executive processes. Given there is no evidence suggesting an ELA for executive tasks, differential effects of EL on Time- and Event-based PM tasks were predicted. This study investigated these predictions. Fourteen participants with neurological memory impairment completed four computer-based PM tasks in a within-subjects 2x2 factorial experiment, with each factor having two levels: encoding method (Errorless, Errorful), and PM task type (Time-based, Event-based). A significant ELA was observed for Event-based PM ($d=.63$), but not for Time-based PM ($d=-.01$), and the interaction between encoding condition and task type approached significance ($d=.41$). Errorless Learning also resulted in reduced accuracy in participants' retrospective estimates of how many opportunities there had been to perform the PM tasks, suggesting that encoding manipulations can affect metacognitive awareness of PM performance. These findings extend the existing evidence for the benefits of Errorless Learning within cognitive rehabilitation, by showing for the first time that EL can benefit future action in addition to performance on purely retrospective learning and retrieval tasks. There are also clear clinical implications of these results; day-to-day Event-based PM tasks (e.g. take your medication with breakfast, check you've got your keys before you go out the front door), if learned with Errorless methods, are more likely to be acted upon than tasks where errors have been made during learning.

2. Introduction

This introduction comprises two main sections, one on Prospective Memory (PM) and one on Errorless Learning (EL). The PM section begins with a working definition of PM and a consideration of the types of task that are thought to use it. It then summarises the available means of measuring PM, and evaluates them from both clinical and empirical perspectives. Next, a theoretical section describes the stages involved in PM task completion and the cognitive and metacognitive processes associated with each of those stages and summarises the dominant theoretical models of PM. It moves on to consider how PM functioning is affected by various neurological conditions. Next, a variety of approaches to rehabilitating PM problems are described, which are broadly separated into task-level approaches that target all major cognitive task demands, and process-level approaches, which specifically target the mnemonic components or attentional/executive components of PM tasks. The EL section reviews the theoretical background, the evidence base for the efficacy of EL in cognitive rehabilitation, and the potential mechanisms that underpin its beneficial effects. Finally, the rationale for the present study of EL in PM rehabilitation is presented, along with a summary of the aims and hypotheses.

2.1. A Working Definition of Prospective Memory

The term “Prospective Memory” (PM) refers to the processes involved in remembering to act upon previously formed intentions. PM tasks are frequently encountered in everyday life; they range from remembering to pay the gas bill or to take medication, to remembering to pick the children up from school. The definition of a PM task is hence quite simple, and the term could encompass any task which involves the formation of an intention that cannot be immediately acted upon. However, it is useful to specify some further parameters. Consider, for example, the task of needing to book a summer holiday, or to turn left at the end of the road. Both tasks have PM components, but they are respectively unique and highly salient, and frequently encountered and overlearned, so it may be more sensible to break them down into smaller sub-goals (e.g. remembering to search for a hotel), or view them as procedural tasks (e.g. as a step in the task of completing a journey from A to B) rather than as PM tasks.

The working definition used here is that a PM task is one that involves the formation of a conscious intention to perform a discrete action or perhaps set of actions, which cannot be immediately acted upon and does not form part of a simple procedural routine. Within Squire and Zola’s (1996) taxonomy of human memory, PM would likely be considered a subtype of episodic memory. The future-oriented nature of PM does superficially distinguish it from other types of memory, and there is some evidence in support of this assertion. For example, memory for future action has a special status relative to memory for past action (cf. the

“intention superiority effect”; Goschke & Kuhl, 1993), and thinking about the future has been shown to place greater demands on frontal/executive functions than thinking about the past (Weiler, Suchan, Koch, Schwarz, & Daum, 2011). Recent research, however, emphasises the great overlap in the brain mechanisms that underlie both prospective and retrospective thought (Schacter, Addis, & Buckner, 2007). Furthermore, there have been no reported dissociations between prospective and retrospective memory, so there is no strong basis for believing that these processes are supported by independent neural systems.

2.2. Systems for Categorising PM tasks

Given the diversity of tasks with a “delayed intention” component, it is understandable that much of the early research on PM involved classifying different types of task, and linking these to everyday memory tasks. Einstein and McDaniel (1996), for example, differentiated between Event-based, Time-based and Activity-based tasks. Within this scheme, Event-based tasks are those that should be enacted in response to some external occurrence, such as posting a letter upon passing a postbox. Time-based tasks are those that should be enacted at a particular time, such as attending a doctor’s appointment at quarter past two. Activity-based tasks are similar to Event-based tasks, but here the triggering ‘event’ is one’s preceding activity, such as switching the oven off after cooking. This distinction is in widespread use in the PM literature.

In one widely used memory questionnaire, Smith et al (2000) distinguish between “self-cued” and “environmentally-cued” tasks, for both prospective and retrospective memory tasks. They equate these categories to the time and Event-based distinction described previously. However, the terminology is not in widespread use, and the distinction is not used in formal scoring procedures for the questionnaire. They also differentiate between “short-term” and “long-term” PM tasks, with the short term questions referring to periods of “a few minutes”, and long term questions referring to appointments, buying birthday cards, and passing on messages, without giving a timeframe. Hence, clear information regarding the boundary between short-term and long-term tasks is lacking, and it is not clear whether the distinction intends to map onto the distinction between working memory and long-term memory, the distinction between immediate and delayed recall within long term memory, or even a more colloquial description. Any of these distinctions may be problematic, as once formed, an intention may be recalled on an intermittent basis a number of times prior to its enactment, and this may serve to standardise the overall delay. This may therefore remove, or at least blur, the distinction between so-called short-term and long-term tasks (Ellis, 1996).

One further subdivision of PM tasks is Ellis’ (1988) distinction between “pulses” and “steps”. This distinction was founded on the basis of diary studies of neurologically healthy volunteers, who recorded their day-by-day intentions, rating them on dimensions such as subjective

importance, and recording whether memory aids were used to remember them. Pulse intentions are defined as those associated with a specific time for action, such as “I must call the bank at 4:00”, or, “I must write that letter as soon as I get off the phone”, whereas step intentions are those with a wider time-frame in which they can be completed, such as “I must call the bank at some point”, or “I must write that letter this week”. Pulses were rated as more personally important than steps, and were more often supported by external memory aids. The pulse/step classification has not caught on in the same way as the time/event/activity one, but other studies have similarly distinguished PM tasks in terms of target specificity, for example by referring to time-specific and non-time specific tasks (Miotto & Morris, 1998), and the concept has an intuitive appeal. Not all of our intentions can be tightly locked to a given event or time interval, and the concept of task specificity that is incorporated in the pulse/step distinction is valuable¹, and something that is rather lacking in the time/event classification.

2.3. The Measurement of Prospective Memory²

PM tasks can span considerable time periods, range from trivial to crucial in importance, and be classified into a number of sub-types. Furthermore, performance on these tasks can break down for a number of reasons. It follows that there are a number of potential approaches to measuring PM. These can be categorised as follows.

2.3.1. Using Standardised Measures of PM

These table-top tests involve an “Ongoing Task”³ to occupy a participant’s attention (e.g. a set of puzzles), alongside occasional Time and Event-based PM tasks. These PM tasks may include remembering to ask a specified question when an alarm sounds, or to switch puzzles at a particular time. Examples include the Cambridge Prospective Memory Test (CAMPROMPT; Wilson, Emslie, Foley, et al., 2005) and Memory for Intentions Screening Test (MIST; Raskin, 2004). PM tasks are also included in all versions of the Rivermead Behavioural Memory Test (RBMT; Wilson, Cockburn, & Baddeley, 1985).

2.3.2. Assessing Component Processes of Prospective Remembering

This approach involves the administration of standardised tests that assess processes thought to contribute to performance on PM tasks. An example battery would include measures of immediate and delayed memory and executive functions, along with potential measures of

¹ It is a well-established that more closely specified intentions are more likely to be acted on than vaguely specified ones (Locke & Latham, 2002), and this principle is emphasised in many goal-setting guides and behaviour change interventions (Hart & Evans, 2006; Michie & Johnston, 2004).

² These approaches were previously described by Fish, Manly and Wilson (2010). However, this section represents an expanded and updated review.

³ A decision was made at the First International Conference on Prospective Memory in 2000 that the term “Ongoing Task” should be used to refer to the non-PM task(s) within the PM paradigm, rather than other alternatives which had previously been in used, such as “cover task” or “background task”.

sustained attention, to get an idea of how able a person may be to form and remember intentions, and to plan and organise for their achievement.

2.3.3. Measuring Performance on “Real” PM Tasks

This approach involves assigning PM tasks that take place outside of the clinic, for example, asking a participant to post a letter to the clinic, or to make phone calls at set times. It is feasible that such tasks could be embedded within an assessment session. For example, when arranging the appointment, one could request that the person completes a form and brings it with them to the session. This could then form a discussion point during the clinical interview, as well as provide some data on everyday PM functioning, in terms of behavioural observations (e.g. are the arrangements for the appointment spontaneously recorded in a diary or calendar, do they request a reminder, do they take up an offer of this reminder, have they filled in and returned the questionnaire, did they need to be reminded to return it, etc.), as well as data amenable to analysis for clinical audit or research purposes. There are several examples of this approach (Fish et al., 2007; Maylor, 1990; Zogg et al., 2010).

2.3.4. Using Self and/or Informant Ratings

There are various published questionnaires that either assess PM, or include PM items. The Prospective and Retrospective Memory Questionnaire (PRMQ; G. Smith et al., 2000); and the Comprehensive Assessment of Prospective Memory (CAPM; Roche, Fleming, & Shum, 2002) are those that have received the most research attention. Both include self and informant-rated scales. Customised PM diaries and/or rating scales could also be used to obtain an estimate of day-to-day PM functioning.

2.3.5. Using Laboratory-Based Tasks Derived from the Literature

These paradigms, like the standardised clinical measures, assign instructions for two tasks, one attentionally demanding Ongoing Task, and one or more PM tasks. The literature includes a great many variations on this theme, the most widely used being computer-based tasks in which the ongoing task is to answer trivia questions, or rate words on various dimensions, and the occasional PM task is to press a designated key when a particular word, e.g. “boat”, appears in the task. More innovative variations on this theme include Virtual Reality-based assessments (Sweeney, Kersel, Morris, Manly, & Evans, 2010) where participants perform more ecologically valid PM tasks (e.g. under the guise of working as a furniture removal person, remembering to check for the arrival of a delivery van, and to label glass items with ‘fragile’ notices).

The advantages and disadvantages of each approach are summarised in Table 1.1, overleaf.

Table 1.1. Summary of the advantages and disadvantages of different approaches to measuring Prospective Memory.

Approach	Advantages	Disadvantages
Standardised tests of PM	<ul style="list-style-type: none"> - Good normative data - Good face validity - Psychometric properties should be adequate (or, at least, information on these should be available) - No additional materials needed - Standardised administration procedures 	<ul style="list-style-type: none"> - Limited evidence of predictive validity - Expensive to purchase - Significant administration time (and cost implications of this) - May not generate much range in performance - Interdependence between components (e.g. failure of the memory component precludes success on the executive component), means may not actually assess all stages of PM, and may be difficult to specify the locus of any problem detected (e.g. whether it stems from loss of intention, distraction, poor initiation, etc.)
“Component process” assessments	<ul style="list-style-type: none"> - There are standardised procedures for the component tests used, and accompanying normative data - Time and resource-efficient as tests would be included in any neuropsychological assessment - Increased flexibility through choice of component tests 	<ul style="list-style-type: none"> - No standardised procedures for compilation of batteries, nor interpretation in relation to PM - Adequate operation of component processes cannot ensure adequate <i>integration</i> of these processes to enable the successful completion of everyday PM tasks
Naturalistic tasks	<ul style="list-style-type: none"> - Good ecological validity - Good face validity - Likely advantage for predictive validity - Potential use to gain insight over deficits and to promote compensation - Can be used to establish effectiveness of compensatory strategies or as more therapeutic ‘behavioural experiments’ 	<ul style="list-style-type: none"> - May be perceived as intrusive/onerous in comparison with a one-off assessment - Hard to establish reasons for task omission/failure - Variation in structure and time commitments of individuals’ day-to-day lives may make it difficult to establish a sound normative basis against which to judge individual performance.
Self or informant-ratings	<ul style="list-style-type: none"> - Good face validity - Quick to complete - PRMQ has good normative data and freely available scoring software - May help identify treatment targets/areas for intervention - Several freely available to use including translated scales 	<ul style="list-style-type: none"> - Remembering to complete a diary is in itself a PM task, and easily forgotten - Retrospective memory demands are high for both questionnaires and diaries (e.g. did I remember my medication? Do I forget people’s names?) - Participant may lack insight into his/her difficulties, or equally be concerned by relatively minor mistakes - Informant ratings may be more reliable but be affected by some of the same factors - In undergraduates, PM self-report is reliable but not valid (Uttl & Kibreab, 2012)
Experimental analogue tasks	<ul style="list-style-type: none"> - Designed to fit with rationale of given experiment - Enhanced control over potential sources of measurement error relative to other methods - Literature provides “tried and tested” approaches, across various groups - Should be able to recreate basic tasks from the papers’ methods sections - May offer preliminary normative data - Many clinical groups have been studied - Potential for more ecologically valid procedures using new technologies such as Virtual Reality 	<ul style="list-style-type: none"> - Uncertain predictive validity - Can be difficult to identify best version to use as there are many published paradigms - May require technical skills/equipment to create tasks - Some tasks may not have been evaluated in clinical populations

2.3.6. The Ecological Validity of PM Assessments

One major criticism of the measurement of complex cognitive operations within a laboratory or clinic setting is that many paradigms fail to sufficiently resemble real-life situations (i.e. they lack "ecological validity"; Wilson, Cockburn, Baddeley, & Hiorns, 1989). In the domain of executive functioning, Burgess et al. (2006) stated that the strong tradition of using experimental tasks in the assessment of executive functioning whilst neglecting more naturalistic measurement had led to the situation that "We know virtually nothing about how the brain allows us to organise simple, everyday activities like cooking, and interacts with environmental factors in doing so [...] research has spent several decades investigating the dynamics of (by comparison) esoteric activities such as performing the Wisconsin Card Sorting Test" (p.197). The concern here seems to be that by boiling down tasks to the level at which they are amenable to measurement within a relatively simple paradigm, there is a danger that the resulting theoretical insights and clinical applications will be limited. Burgess et al. (2006) do argue that if there is thoughtful consideration of the concordance between the operations likely required within the experimental task and those required within the everyday situation they are intended to mimic, however, then these limitations may be less apparent. There are additionally clinically-motivated concerns about the use of these relatively esoteric tasks; for example they can be aversive for patients to complete, the results can be difficult for non-specialists to understand, and the patient/family may feel that their everyday difficulties have not been understood.

Modern published clinical tests do include validity studies, but the evidence is often rather weak, including (as noted in Table 1.1) the evidence for standard clinical tests of PM. One factor that potentially contributes to this is that it is difficult to reliably measure 'everyday functioning'. One notable exception is the original version of the RBMT (Wilson et al., 1985), scores on which had a strong correlation with scores from extensive behavioural observations of memory failures (an average of 35 hours per patient; Wilson et al., 1989). Even more compellingly, RBMT scores during rehabilitation differentiated between patients categorised as independent versus dependent at 10-year follow up. Current RBMT scores had even stronger discriminatory power, whereas current scores from the Wechsler Memory Scales had no such discriminatory power (Wilson, 1991). Nonetheless, the specific validity of the PM items from the RBMT, and its revised editions, remains unknown.

It is also somewhat unclear whether or there may be differences in reliability between ecologically valid and ecologically less valid tests of PM. Take Shallice and Burgess' (1991) Multiple Errands Task, for example, which uses shopping as a model everyday activity that taxes executive capacities. One might predict that as most people will already be familiar with

the task of having to shop for various items within a restricted time period, novelty factors would have a relatively lesser effect on this ecologically valid task than a traditional laboratory-based equivalent. This suggests that test-retest reliability may be relatively better for the ecologically valid measures. This may in turn mean that similar versions of the same task may produce similar types of performance, which may then suggest that superior parallel form reliability would also be expected. However, as behaviour in these tasks is less tightly constrained and may be more difficult to directly observe, it may be more difficult to accurately rate performance, and as such inter-rater reliability may suffer relative to more traditional tasks. Additionally, given the fact that ecologically valid tests take place in more life-like scenarios, participants are likely to be more vulnerable to distraction in these scenarios than in a laboratory setting. Hence it is possible that performance may vary more from one occasion to the next, with a concomitant decrease in reliability relative to more constrained paradigms. It is also likely that such measures would better predict everyday shopping ability than other everyday executive tasks such as those involved in meal preparation, or running a household. The same arguments seem likely to apply to tests of PM.

Though the ecologically valid measures are important, especially in certain circumstances (e.g. where they are more acceptable for the patients completing them), in the domain of PM we do not yet know whether there is an advantage for clinical tests over experimental measures in ascertaining a given patient's likely everyday functioning. Thinking specifically about brief computerised measures of PM, they do obviously simplify the response properties in relation to everyday life (e.g. pressing a button instead of posting a letter), the time scale may be rather briefer than everyday tasks, and the Ongoing Task is likely much more structured and invariant than everyday distractions. However, in their favour, they have a "dual-task" nature that reflects many everyday PM situations, and crucially they involve the encoding, retrieval and enactment of intentions as do everyday PM tasks. Further, the time-scale difference may be reduced by the intermittent retrieval that is found in everyday PM. Of course, these tasks also allow for accurate and efficient data collection. It is important, therefore, not to reject the use of computerised tasks simply because they lack some of the more dynamic aspects of everyday PM tasks.

2.3.7. Recommendations for Measuring PM

In conclusion, though tasks developed for the experimental analysis of PM in healthy participant groups may have clinical use, this has not been firmly established. There are some merits in taking a componential approach and inferring likely barriers to successful PM functioning, although if the PM problem comes from difficulty combining component skills, this approach may be insensitive. Giving participants "actual" PM tasks to complete is

potentially an informative exercise, although, depending on the specific properties of the paradigm, the reasons for PM failure may not be adequately specified. It is also important to consider general issues such as the impact of strategy use on PM performance, in both everyday life and the chosen assessment options. For example, in the CAMPROMPT, participants are permitted to take notes on the PM tasks to be remembered. This may greatly enhance performance on the test. However, people who take notes in the CAMPROMPT, when the option is expressly offered, may not take notes in everyday life. Conversely, if a test prevents the person from using a strategy that they may use in everyday life, then this would not necessarily produce an accurate representation of their likely everyday PM performance.

Another issue important for measurement is the degree to which a person has insight into their PM performance. Roche, Fleming and Shum (2002) found that people with TBI underestimated how often they made PM errors relative to close informants, indicating reduced insight (though of course the assumption that informant ratings are veridical is not without problems). Moreover, Uttl and Kibreab (2011) have found that in a sample of University students, though self-report ratings of PM had acceptable reliability, they had no association with behavioural measures of PM performance, and hence may not be a valid form of measurement. As a variety of questionnaires were used in this study, it seems more likely that this results from a problem with using self-report methods in this domain, rather than a problem with the particular self-report measures available in this domain.

In clinical settings, then, it would appear preferable to combine questionnaires and component-part measurement with a careful interview regarding everyday PM performance. For many rehabilitation procedures, naturalistic tasks are likely to be the best outcome measures, though questionnaires may give valuable additional information. Laboratory measures do not necessarily have great application within day-to-day clinical work, as evidence regarding predictive validity and reliability is lacking. However, the greater efficiency with which data may be collected in comparison with naturalistic measures means that computerised tasks offer a pragmatic option for use in experimental investigations whether those tasks take place in a virtual reality or less naturalistic paradigm.

2.4. Theoretical Perspectives on Prospective Memory

This theoretical section aims to outline the hypothesised stages of PM, and to explore the potential cognitive contributions to each of those stages, to examine specific models of retrieval in PM, and to review the neuropsychology and broader neuroscience of PM.

2.4.1. Hypothesised Stages of PM

Ellis (1996) outlined a number of stages involved in completing PM tasks, beginning with the encoding of task-relevant information, the retention of that information over a delay, retrieval of that information during the appropriate performance interval (and potentially also outside of the appropriate performance interval), and subsequent action and evaluation of performance. These stages may take place over sometimes extended time periods, and in the context of other, often un-related, activities.

2.4.2. Cognitive Requirements of PM Task Completion

Fish, Manly & Wilson (2010) emphasised the hierarchical nature of PM tasks, whereby memory for the intention is a prerequisite for successful task completion, yet insufficient to ensure completion. Later stages of PM tasks require not only that the memory for the content of the intention remains intact, but also that attentional and executive processes (i.e. those processes that guide behaviour towards goals; Kopp, 2012) are engaged to allow the retrieval cue (be it a particular time or event) to be noticed, the intention to be retrieved and then acted upon, and all this in the context of many other potential demands including concurrent activities that may distract from the inactive intention. There are also metacognitive aspects of PM, including task-specific awareness of errors and evaluation of performance, and broader insight into one's general PM abilities. A schematic illustration of the stages involved in prospective remembering and their cognitive components is provided in Figure 1.1.

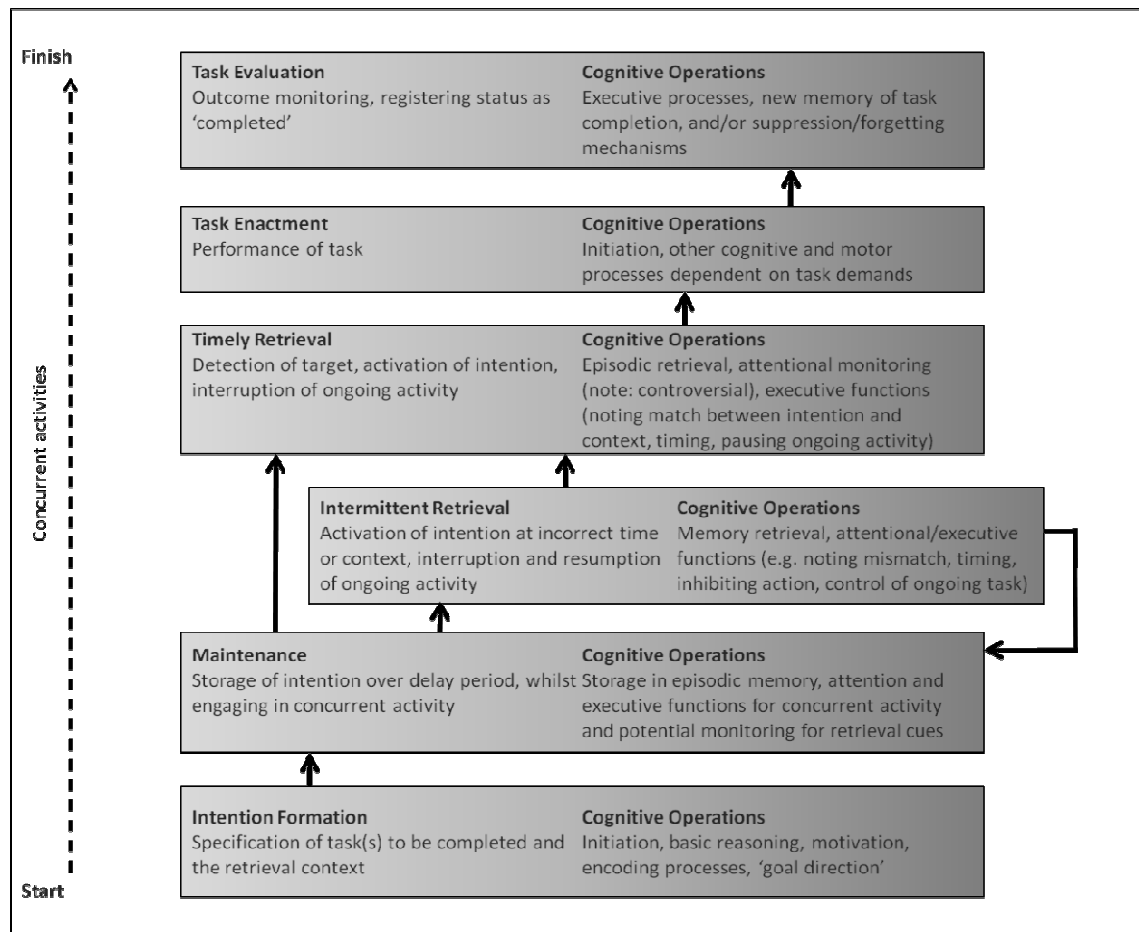


Figure 1.1. Schematic diagram showing stages involved in completion of Prospective Memory tasks and associated cognitive operations.

Note that the precise processes involved are likely to vary according to the demands of individual PM tasks, and failure at any stage prior to enactment may lead to PM task failure. Note also that for clarity, further processes such as updating and reforming intentions are not shown, but could occur at any point and would simply involve a return to the initial intention formation stage.

2.4.3. Models Specifically Examining Retrieval Processes in PM

As should be clear from Figure 1.1, retrieval is only one of several processes important for the completion of PM tasks. However, models of retrieval in PM are dominant within the literature, and will hence be the focus of this section. These models are the Multiprocess Model, and the Preparatory Attention and Memory Model.

2.4.3.1. The Multiprocess and Preparatory Attention and Memory Models

Einstein, McDaniel and colleagues (Einstein & McDaniel, 1996; Einstein et al., 2005) have posited an influential Multiprocess Model, which states that there can be multiple processes underlying PM retrieval. They differentiate between *strategic* processes, including rehearsal of the intention and monitoring of the environment, and *automatic* processes, related to the

reflexive triggering of the stored intention in response to an environmental cue. These automatic and strategic processes are thought to depend upon medial temporal and frontal lobe functions respectively (McDaniel & Einstein, 2000). There is considerable debate within the literature, however, regarding whether it is *ever* possible for PM retrieval to take place without the involvement of strategic processes. The main proponent of this view is Smith (2003), who has proposed a Preparatory Attention and Memory Model, which states that some attentional or strategic resources are always necessary to monitoring the environment and hence allow PM retrieval to take place.

Numerous studies published by the McDaniel-Einstein groups (e.g. Einstein et al., 2005; Scullin, McDaniel, & Einstein, 2010) have reported that, within the constrained laboratory tasks previously described, PM retrieval can take place without any detectable detrimental impact on ongoing task performance in terms of accuracy or response latencies. This is taken as evidence that ‘automatic’ retrieval has taken place, as the use of ‘strategic’ processes would logically have a detrimental effect on ongoing task performance. Harrison and Einstein (2010), for example, found no cost to ongoing tasks when the PM task was for the participant to press a key when their own name appeared within the task. Smith (2010), however, reviewed the studies purporting to demonstrate automatic PM retrieval and concluded that the experiments simply lacked statistical power to detect costs to the ongoing task, and all studies in fact showed the trend predicted by the PAM model. It appears from these experiments that PM retrieval can take place with very little strategic involvement, but this does not mean that such retrieval is entirely lacking an attentional component. Indeed, Smith (2010) noted that Einstein and McDaniel (Einstein & McDaniel, 2010) had moderated their terminology, stating that the term “spontaneous retrieval” was preferable to their previous “automatic retrieval”. If such processes are not viewed as automatic then there is little to distinguish the two models.

Looking at the example of intending to post a letter, it is clear that seeing a letterbox constitutes a strong retrieval cue to trigger the intention. This might indicate that the task requires minimal strategic processing. However, the fact that one passes letterboxes every day may reduce their environmental salience and necessitate the involvement of more active monitoring processes to spot their relevance for the stored intention to post the letter. It therefore seems likely that real-world PM tasks lie on a continuum involving strategic and mnemonic processes to support retrieval to varying degrees, depending upon the particular characteristics of the task at hand. Although there is no clear-cut dissociation in terms of the cognitive processes supporting time versus Event-based PM, there is evidence that Time-based tasks rely to a greater extent upon strategic processes than more automatic retrieval, and

Event-based tasks the reverse, as evidenced by studies reporting disproportionate age-related decline in Time-based relative to Event-based PM (d' Ydewalle, Bouckaert, & Brunfaut, 2001).

2.4.3.2. Factors that Modulate PM Retrieval

Many studies of PM show that timely PM retrieval can be modulated by factors specific to the properties of the intended action. PM performance is improved under the following conditions relative to comparison conditions (note that this is an illustrative rather than exhaustive list):

- when the PM target is particularly distinctive or familiar, or directly in the focus of attention (Brandimonte & Passolunghi, 1994; Harrison & Einstein, 2010; Kliegel, Jäger, & Phillips, 2008)
- when there is a semantic association between the retrieval cue and intended action (Nowinski & Dismukes, 2005)
- when the PM task is perceived as more important than the ongoing task (Alexander, Martin, McDaniel, & Einstein, 2004; Kliegel, Martin, McDaniel, & Einstein, 2001)
- when there is an incentive associated with the PM task (Meacham & Singer, 1977)
- when there is a briefer delay between encoding and retrieval (Morgan, Weber, Rooney, Grant, & Woods, 2012)
- when a specific “when-then” statement outlining task requirements has been formed (e.g. when the word “apple” appears I will press the space bar) compared with a control statement (e.g. I want to press the space bar when the word “apple” appears) (McDaniel & Scullin, 2010)

Importantly, such modulation is not dependent on outright ‘forgetting’ of the intention in comparison conditions. That is, when asked at the end of the test directly about the tasks they were to complete, all participants remember what they were supposed to do and when they were supposed to do it. However, they are more likely to *act upon* that intention with the specified manipulation. This implies that properties of the intention itself can impact upon retrieval. There are important applications of these findings, as they suggest that when the salience, importance and distinctiveness of PM task content is increased, successful task completion is more likely to result. Therefore, artificial or effortful enhancement of these aspects of PM tasks may be an effective strategy for people who wish to improve their PM performance. Indeed, simple strategies to support PM do employ such principles. For example, a common technique to prevent forgetting an important but non-routine belonging when you leave the house is to place that object directly in one’s way. This could be understood as making the to-be-remembered object a focal rather than distal PM target.

2.4.4. Metacognitive Aspects of Prospective Memory

Very little research has been conducted into metacognitive aspects of PM. One could easily imagine however, various metacognitive processes or judgements related to PM. For example in lower-level online recognition of one's errors and correct responses as and when they occur, in retrospective evaluation of task performance, and in higher-level beliefs about one's general abilities in this domain. Agnew and Morris' (1998) model of awareness of memory deficits in Alzheimer's disease is informative in this regard. The model differentiates between three types of anosognosia. Executive anosognosia results from failure at the level of comparator mechanisms within the central executive system. Though errors are perceived, no comparisons are made with past performance and as such insight cannot develop. Mnemonic anosognosia results from failure to encode the mismatch between past and current memory performance in semantic memory. Such patients may be expected to have implicit but not explicit knowledge of their impairments. Primary anosognosia would be predicted to result from a failure to recognise the implicit error signal as an error signal, and though implicit knowledge may remain as with the mnemonic version, there is a different underpinning cognitive mechanism. It is clear that the model could apply to awareness of PM ability also.

Task-level metamemory has been investigated in a small number of studies by asking participants to estimate their level of accuracy on PM tasks before and after completing the PM task (these judgements are referred to as predictions and "postdictions" respectively)⁴. Meeks, Hicks and Marsh (2007) found that within undergraduate samples, such judgements had significant correlations with actual PM performance, however the correlation was only weak for predictions ($r=.23$), and moderate for postdictions ($r=.63$), suggesting that they are not very accurate. Smith, Souchay and Moulin (2011) found that relative to controls, people with Parkinson's disease made inaccurate predictions of their PM performance, but in contrast, their "postdictions" did not differ from controls.

Another related avenue of research has been on the "output monitoring" stage of prospective remembering. Ellis and Freeman (2008) note that this has not been the focus of much research. They note, however, that failures at this stage of a task could result from either holding a mistaken belief that one has performed a task one has not, or that one has omitted a task that one has actually performed (both metacognitive errors). The former belief would lead to omission errors, and the latter to commission errors. Harvey and Ellis (2005, cited in

⁴ It is interesting to note that within undergraduate samples, the mere act of making performance predictions has been found to improve PM performance (Meier, von Wartburg, Matter, Rothen, & Reber, 2011).

Ellis & Freeman, 2008) found that older adults made more such errors of commission than younger adults, particularly when performing ongoing tasks that had a high attentional load.

Two questionnaire studies in clinical samples have examined potentially higher-level insight into PM difficulties by obtaining discrepancy scores from self and informant ratings of PM performance. These studies have identified reduced awareness of PM problems, for example in people with Schizophrenia (Chan et al., 2008), and in people with ABI (Roche et al., 2002). However, the assessment of PM using only questionnaire measures is rather problematic as it considers the informant ratings to be ‘true’. This is particularly important given that a recent study found no correlation between subjective and objective measures of PM in an undergraduate sample (Uttl & Kibreab, 2011).

An interim conclusion is therefore that, just as in other domains of cognition, PM-related awareness is a multifaceted concept. It may be difficult to accurately assess, and mere assessment may actually interfere with other aspects of the task (e.g. by causing participants to prioritise the PM task over the ongoing task). However, despite the relatively limited body of research into this matter, there are potentially important implications. In particular, this research indicates that self-report measures alone are not a sufficient method of assessing PM, or even for assessing awareness of everyday PM functioning.

2.4.5. Neuroscience Perspectives on PM

Many accounts of the behaviours associated with frontal lobe or “executive” dysfunction include descriptions of problems acting on intentions, in the context of well-preserved abilities in the domains of language, memory and general intellectual functioning (Duncan, Emslie, Williams, Johnson, & Freer, 1996; Eslinger & Damasio, 1985; Luria, 1966; Shallice & Burgess, 1991). Shallice and Burgess’ (1996) proposed “Supervisory Attention System”, which attempted to explain how the cognitive processes of the frontal lobes support complex goal-directed behaviour, includes a hypothesised process labelled ‘delayed intention marker realisation’. This can be seen as synonymous with PM retrieval processes. A report by Cockburn (1995), however, was the first to associate the cognitive errors of a patient with bilateral frontal lobe damage as PM failures, caused by putative “dysexecutive” impairments of planning, initiation and inhibition.

A series of studies by Petrides (1985, 1997) demonstrated that patients with unilateral frontal lobe lesions show impaired learning of conditional associations. That is, in forming links between particular stimuli and particular responses, which is a concept that has considerable overlap with PM task requirements (i.e. in the conditional association “perform action X in the presence of stimulus 1, and action Y in the presence of stimulus 2”, there is an intention for

later action that is dependent on the occurrence of some specified conditions). Petrides' studies also found that patients with unilateral temporal lobe lesions that encompass the hippocampus show equivalent though material-specific impairments, whereby those with right temporal lesions are impaired only for spatial tasks, and those with left temporal lesions only for nonspatial tasks (Petrides, 1985, 1997). PM tasks can be distinguished from conditional associative learning tasks by the ratio of PM task to ongoing task trials, PM trials being rare within a PM paradigm, whereas the two types of trial in the Petrides paradigm would occur with equal probability. Nonetheless, these studies make clear that frontal lobe damage and consequent disruption of certain executive and mnemonic functions may well contribute to PM failures in everyday life.

Looking specifically at PM, detailed neuroanatomical and neuropsychological analyses both in normal ageing (Gordon, Shelton, Bugg, McDaniel, & Head, 2011; McDaniel, Glisky, Rubin, Guynn, & Routhieaux, 1999) and in people with focal lesions following tumour resection (Burgess, Veitch, de Lacy Costello, & Shallice, 2000; Volle, Gonen-Yaacovi, de Lacy Costello, Gilbert, & Burgess, 2011) report consistent findings that damage to temporal lobe structures, particularly the hippocampus, is associated with poor performance on both Event-based PM tasks and tasks that require minimal monitoring in order to detect targets (Gordon et al., 2011). Performance on tasks that require substantial monitoring for successful completion (e.g. Time-based tasks, but potentially also Event-based tasks in which targets are difficult to detect), however, is adversely affected by frontal lesions, more specifically to right frontopolar cortex (Volle et al., 2011).

Burgess et al.'s (2011) recent review of the burgeoning functional imaging literature reports a remarkable consistency of results. Performing PM tasks, or merely holding an intention for action in mind, is associated with increased Blood Oxygen Level Dependent (BOLD) signal activations within *rostrolateral* prefrontal cortex relative to control tasks. Further, performance on control tasks is associated with increased *rostromedial* prefrontal cortex activation relative to the same task with a PM component. In addition to this rostral-medial dissociation within prefrontal cortex, co-activations in frontoparietal areas (anterior cingulate, precuneus and wider parietal lobe) are also frequently reported. This latter pattern, however, is far from unique to PM tasks. Indeed, it seems likely to reflect the documented pattern of common regions of frontoparietal cortex recruited in the completion of diverse cognitive tasks (Duncan & Owen, 2000). Burgess et al. note that most of the reviewed studies focus only on Event-based PM tasks. The one study that contrasted Time-based and Event-based PM found increased activation in areas of rostral prefrontal areas, for Time-based relative to Event-based tasks (Okuda et al., 2007). This is partially consistent with the neuropsychological results

previously outlined. Given the neuropsychological data, one would predict that the reverse contrast (event minus time) might show increased activations within areas of the medial temporal lobe⁵. However, this was not so. Rather, a different region of left superior frontal gyrus was identified. Further research on this matter is clearly necessary to clarify these issues; however, what is currently clear is that frontal or executive contributions to PM are important and that damage to prefrontal cortex will have a likely impact on PM function.

2.5. Prospective Memory in Neurological Disorders⁶

Considering the ubiquity of PM tasks in everyday life, and the demands they place upon a variety of cognitive domains, it is not surprising that PM failures are familiar to many. Indeed Baddeley (1997) stated that when normal laypeople talk about having a ‘bad memory’ they are often referring to their PM. For people with cognitive impairment, however, PM tasks present a more particular problem, due to their aforementioned reliance upon mnemonic, attentional and executive functions, which are frequently compromised in brain injury of various aetiologies (Teasdale et al., 1997). Indeed, Kinsella et al. (1996) noted that failures of PM were the *most* frequently reported cognitive problem by a group of people with Traumatic Brain Injury (and as the frontal and temporal cortices are particular vulnerable in TBI this is again not too surprising).

There is considerable evidence that PM performance is impaired in a wide range of disorders including *early dementia* (Huppert, Johnson, & Nickson, 2000), *acquired traumatic and non-traumatic brain injury* (Brooks, Campsie, Symington, Beattie, & McKinlay, 1987; Groot, Wilson, Evans, & Watson, 2002; Schmitter-Edgecombe & Wright, 2004; Shum, Valentine, & Cutmore, 1999), *Parkinson’s disease* (Katai, Maruyama, Hashimoto, & Ikeda, 2003; Kliegel, Phillips, Lemke, & Kopp, 2005), *schizophrenia* (Elvevag, Maylor, & Gilbert, 2003; Kondel, 2002; Shum, Ungvari, Tang, & Leung, 2004), *HIV-associated neurocognitive disorders* (Carey, Woods, Rippeth, Heaton, & Grant, 2006), *substance use disorders* (Weinborn et al., 2013) *bipolar disorder* (Chan et al., 2012), *Attention Deficit Hyperactivity Disorder (ADHD)* in both children (Kerns & Price, 2001; Zinke et al., 2010) and adults (Fuermaier et al., 2013), children with *Autism Spectrum Disorders* (Brandimonte, Filippello, Coluccia, Altgassen, & Kliegel, 2011), and people with *hippocampal sclerosis* in the context of epilepsy (Adda, Castro, Além-Mar e Silva, de Manreza, & Kashiara, 2008). PM deficits have even been reported in so-called *sub-clinical compulsive checkers* (Cuttler & Graf, 2009).

⁵ As all tasks have mnemonic demands this may not be a strong prediction.

⁶ Note that though the research reviewed in this section clearly contributes to the neuroscientific understanding of PM outlined in the previous section, it is presented separately in light of its particular relevance for clinical work.

Within these studies, it has not always been possible to identify the locus of PM impairments. In group studies there would be considerable inter and intra-individual variability, so it is unlikely that one could specify the true locus of every PM task failure for every participant. However, some degree of specification is possible. Studies reporting deficits on simple laboratory paradigms (e.g. Chan et al., 2012), for example, have indicated that the relative impairment is within the stages of noticing the PM target, retrieving the intention to act on it, and initiating the associated action (i.e. the timely retrieval and task performance boxes of Figure 1.1, p. 15). Failure at the intention formation phase is ruled out as the experimenter stipulates the instruction, and storage/maintenance failures are ruled out through testing retrospective memory at the end of the experiment. In other studies, where more naturalistic or multi-componential measures of PM have been used, it is even more difficult to determine the precise impairment(s). Some more complex PM tasks do allow for the generation of separate scores for different components of PM task completion, and when these measures have been used, (e.g., Altgassen & Kliegel, 2011; Kliegel et al., 2005) apparently more clearly circumscribed deficits can be identified. For example, deficits in planning have been associated with ADHD (Altgassen & Kliegel, 2011) and deficits in intention formation in Parkinson's Disease (Kliegel et al., 2005). However, given the aforementioned hierarchical structure of PM tasks, when impairments are detected at relatively early stages of PM, it remains unclear whether or not there are additional impairments at subsequent stages. Despite this degree of confusion regarding the precise nature of PM impairments within these clinical groups, the presence of such difficulties and their significance for everyday functioning and as a target for rehabilitation, are abundantly clear.

2.6. Rehabilitation of Prospective Memory

The consequences of PM problems can be many and varied. There are obvious practical and social repercussions of forgetting particular intentions, for example incurring penalty fees from missing a payment, needing more extensive dental treatment after missing appointments, or offending friends by forgetting to send a birthday card. In addition, there can be important psychological ramifications of repeated PM failure. For example, negative thoughts and feelings of guilt, shame and frustration about PM failures could impede processes of adjustment to and effective coping with cognitive impairment (processes that predict outcome in brain injury; Doering, Conrad, Rief, & Exner, 2011; Kervick & Kaemingk, 2005). The rehabilitation of PM is, therefore, an important target within individual cognitive rehabilitation, within broader goal-focussed neuropsychological rehabilitation and even within psychotherapeutic work, where PM may be addressed as an area of skill development, or as an

implicit target when working on unhelpful appraisals or in supporting completion of out-of-session therapy tasks.

The rehabilitation of cognitive deficits generally takes one of two forms: attempts to restore the deficient functions, or attempts to compensate for the deficits. Cognitive rehabilitation as it is clinically practiced tends to focus on the compensatory approach, as there is little evidence that impaired cognitive functions can be effectively restored (Wilson, 2009).

2.6.1. Restorative Approaches

The restorative approach to rehabilitating PM has often taken the form of practising PM tasks. There are reports from 5 single cases (Sohlberg & Raskin, 1996; Sohlberg, White, Evans, & Mateer, 1992a, 1992b) involving extensive repeated practice of simple PM tasks (e.g. "Raise your hand when I ring this bell") over increasing time delays. Significant improvements were evident on the trained tasks, but very little generalisation was observed. These disappointing findings and the rather arduous character of the training program, has meant that there has been little apparent take-up of these techniques in clinical practice.

There is more recent evidence that repetitive training of working memory (WM) results in benefits that generalise from training tasks to tests of related cognitive functions, and even everyday functioning. As WM contributes to PM performance, the findings are particularly relevant. The most encouraging evidence to date comes from studies of children with ADHD, where benefits from WM training have been shown to generalise to tests of executive functions and fluid intelligence, and to ratings of everyday attentional behaviour (Holmes, Gathercole, & Dunning, 2009; Klingberg et al., 2005; Melby-Lervåg & Hulme, 2013). There is also some preliminary evidence of similar results within adults with acquired brain injury, but it is of a relatively poorer methodological quality than the developmental studies cited (e.g. a study by Westerberg et al., 2007 showed beneficial effects of training but relative only to a waitlist control group). As WM may be seen as a process contributing to PM performance, it is possible that carry-over effects from WM training to PM task performance may be seen.

2.6.2. Compensatory Approaches

In contrast with the above, compensatory approaches are used extensively in rehabilitation for cognitive deficits, and the strategies are frequently directly aimed at improving everyday functioning, so the lack of generalisation is not of primary concern (i.e. that tasks/situations in which the problems are most apparent are those targeted in rehabilitation). Compensatory strategies can range from simple strategies to aid memory such as carrying a note-book, to the elaborate environmental adaptations found in "smart homes" including appliances that use tracking technology to switch off appliances when a person leaves the house, and systems to

automatically detect wandering behaviour and in response provide suggestions for alternative activities (e.g. Jasiewicz et al., 2011; Storey, 2010). Within PM rehabilitation, compensatory strategies can be broadly divided into those that address some or all of the cognitive components involved in task completion (referred to here as “task-level strategies”), and those that address specific cognitive components of task completion (referred to here as “process-level strategies”, and subdivided into executive and mnemonic strategies).

2.6.2.1. Task-Level Strategies for PM

Several studies have found that electronic memory aids, which remind the person of the intended task at the appropriate time, are effective at increasing the rate of PM task completion. The best examples of such studies are those of the NeuroPage system, which uses a central computer to send reminder messages to simple pagers worn by the person with memory impairment (Wilson, Emslie, Quirk, Evans, & Watson, 2005; Wilson, Evans, Emslie, & Malinek, 1997). In the 2005 study, 85% of the 143 participants were found to have significantly benefitted from using the paging system. Even people with severe memory problems can learn to use these devices, particularly if caregivers are involved in training procedures (e.g. in learning to use the device, contacting administrators to update messages and so on). More complex memory aids that have been evaluated in smaller-scale though similar projects include mobile phones (Kim, Burke, Dowds, Boone, & Park, 2000), and smartphones (Svoboda, Richards, Leach, & Mertens, 2012).

A further approach to PM rehabilitation has been to develop programs involving psychoeducation and training in a variety of compensatory strategies to support PM functioning, and there are promising findings from small-scale studies (Kinsella et al., 2009; Shum, Fleming, Gill, Gullo, & Strong, 2011).

In pragmatic terms it is important for people to try out various strategies and identify those that work best for them, but it is of course equally important to isolate effective components of therapy through research, and to identify specific effective techniques, to allow for a targeted approach to PM rehabilitation. These may be best examined through experimental studies of strategies addressing specific task components.

2.6.2.2. Process-Level Strategies Targeting Executive Demands

There are few studies examining strategies that target the executive component of PM tasks. Given the clear executive demands of PM tasks this may be seen as surprising; however, it may simply reflect the rather esoteric nature of research on PM, and natural clinical inclinations to maximise the probability of success by using strategies that address as many aspects of the

target tasks as possible (e.g. by addressing mnemonic, executive, practical and emotional aspects of the tasks rather than a single component).

Recent studies have demonstrated that supporting the ‘monitoring’ component of PM tasks can result in significant improvements in performance. Fish et al. (2007), for example, trained people with Acquired Brain Injury (ABI) in a very brief form of Goal Management Training (Levine, Robertson, & Manly, in preparation) which essentially included psychoeducation about the nature of PM after brain injury, and the description of a simple strategy: to take occasional pauses in activity (or inactivity) to briefly review one’s current and future goals, and to make changes/act upon those goals as necessary. This strategy was associated with a mnemonic - “STOP!”- denoting Stop, Think, Organise and Plan. The aim was that engaging in such reviews would increase participants’ monitoring and sense of ‘goal direction’. PM performance was measured over a two-week period with the task of making phone calls at set times, and use of the goal management strategy was supported by sending occasional text messages containing the “STOP” mnemonic, on a random selection of days, and importantly at times that were not within the 30-minute time window prior to the PM call-times. Hence, the messages did *not* simply prompt immediate action in the way that the NeuroPage system does. PM performance was found to be far superior on days where “STOP” messages were sent compared with days without those messages, which indicates that supporting the executive component of PM tasks has significant potential as a rehabilitation strategy. This effect has recently been replicated in a paediatric sample (Rous, Adams, Fish, Manly, & Adlam, 2012), and extended into an RCT format, where beneficial effects were found to generalise from the phone call task to participants’ achievement of individually specified goals (Manly et al., in prep).

2.6.2.3. Process-Level Strategies Targeting Mnemonic Demands

Several studies have examined the impact of encoding strategies on PM performance, which address the mnemonic component of PM tasks. Specifically this has included the impact of spaced retrieval, elaborated encoding, and mental imagery.

Spaced retrieval (also referred to as expanded rehearsal) is a term that refers to the repeated recall of to-be-learned information over gradually increasing time intervals (e.g. initially after no delay, then 10s, 30s, 2mins, 5mins etc). This is an established means of improving recall (Bjork, 1988; Landauer & Bjork, 1978). McKittrick, Camp and Black (1992) found that learning the content of a PM task (to check a noticeboard to identify one’s daily tasks), with spaced retrieval led to improved performance of the PM task relative to baseline in four people with Alzheimer’s Disease (AD). Two further controlled studies have found that the same method leads to improved PM performance in people with early AD relative to a simple repetition

control learning procedure, and that elaborated encoding (simply practising the PM task prior to the spaced retrieval condition), further improved subsequent PM performance (Kinsella, Ong, Storey, Wallace, & Hester, 2007; Ozgis, Rendell, & Henry, 2009).

Though the results of these studies are important and certainly have clear implications for clinical practice, it is as yet unclear whether PM performance in these studies is facilitated by spaced retrieval simply because the content of the intention is retained in one condition and lost in the other. That is, we do not know whether this encoding technique simply prevents a failure at the low-level mnemonic stages of the hierarchical model; or whether there might be additional benefits at the higher-level stages involving retrieval and action initiation components. The cognitive psychological literature would suggest this to be the case, but direct evidence of this has been lacking. One recent study has, however, reported an effect that hints at such higher-level benefits. Grilli and McFarland (2011) compared a self-imagery encoding strategy on the laboratory-based PM performance of people with non-progressive memory impairment with a control rote rehearsal condition (i.e. participants imagined themselves performing the task, or verbally repeated the task instructions for 45s). PM performance was superior in the imagery condition, and importantly, this was in the context of participants' post-test recall of task demands being intact in both experimental and control conditions. This suggests that encoding strategies may have benefits that extend beyond the initial stages of intention formation, to the timely retrieval and enactment stages of PM tasks.

2.7. Errorless Learning in Memory Rehabilitation

2.7.1. Origins of Errorless Learning

The term "Errorless Learning" (EL) was first used by Terrace (1963a) in animal studies of discrimination learning. Before this time, it was thought that errors must be made and then either punished or extinguished in order for learning to take place. Terrace, however, found that pigeons could be trained to discriminate between two colours without making errors during learning. This was achieved by initially exposing the pigeons to only the 'correct' colour, and positively reinforcing responses to it. Later, the incorrect stimulus was introduced, but only for such brief durations that it effectively removed the opportunity for the pigeons to make errors. Over time, the incorrect stimulus could be presented for long durations without the pigeons responding to it, showing that they had learned to respond to the correct stimulus only, without any punishment or extinction of erroneous responses.

EL was associated with an absence of the negative by-products associated with traditional discrimination learning (i.e. peak shift, behavioural contrast⁷). Additionally, there were benefits including an increased resistance to extinction, reduced aggressive responding to non-reinforced stimuli (Terrace, 1963a), improved inhibition to non-reinforced stimuli (Terrace, 1963b), and facilitation of transfer to a new task when that new task was faded in with the established (Terrace, 1963c). Terrace (2010) noted that experimental interest in EL declined after studies (e.g. Kodera & Rilling, 1976) showed that it was actually associated with the previously mentioned negative by-products of discrimination learning, albeit to a much lesser degree. Clare and Jones (2008) note that though these studies are very often cited as the theoretical basis for using EL in clinical practice, clinical studies frequently ensure errorless performance by means of Skinnerian techniques such as chaining and shaping, which have a longer history.

2.7.2. Techniques for Errorless Learning

Wilson (2010) defines EL as “a teaching technique whereby people are prevented, as far as possible, from making mistakes while they are learning a new skill or acquiring new information” (p. 89). One canonical instruction used during EL is to remind the learner to avoid guessing; that if they are not sure then you will provide a clue. Beyond this instruction, there are a number of ways of minimising errors (described in Clare & Jones, 2008; Wilson, Baddeley, Evans, & Shiel, 1994; Wilson, 2009). These include:

- The provision of verbal or written instructions
- The use of “vanishing cues”, that is, initially providing full information and asking for it to be repeated, then following successful repetition, progressively reducing the information given and asking the person to repeated recall the full information (but only after previous successful recall, and if they’re sure they know the answer).
- Broader “chaining” approaches. Chaining methods involve breaking a task down into component steps, and forward chaining refers to building up competence from initial performance only of step 1, to step 1 and 2, then 1-3, etc. Backward chaining involves demonstration of the whole task, then demonstration of all except the final step which the learner is then to perform, then same procedure gradually increasing the learner’s steps so that the task can be completed independently.

⁷ The “peak shift effect” refers to the increase in the conditioned behavioural response observed when exaggerated versions of the conditional stimulus are presented, and “behavioural contrast” refers to a change in the conditioned response to one stimulus evoked by an alteration to the reinforcement properties of another, independent, stimulus.

- Spaced retrieval, where information is rehearsed over increasing time intervals. Though this is a rehabilitation technique in its own right, it can be used in Errorless Learning if the time intervals are spaced such that errors/failures are less likely.

There has been some confusion or controversy over what constitutes EL. Some authors have thought that this necessitates a didactic approach with little active participation from the learner (e.g. through repeated presentation of information; Hodder & Haslam, 2006), but this need not be the case. Indeed, Wilson et al. (1994) expressly stated that active participation is likely to be very important. An Errorless Learning approach simply involves doing whatever is necessary to “scaffold” a task, thereby minimising incorrect responses and maximising the likelihood of correct responses.

2.7.3. Applications of Errorless Learning Methods

Errorless Learning techniques have been applied in a variety of populations. The first studies were in the field of intellectual disability, where EL was found to be superior to trial-and-error learning in simple discriminations tasks (akin to those used in Terrace’s experiments), as well as more practical tasks such as the identification of coins and letters (Jones & Eayrs, 1992). EL has also been shown to be more effective than control learning procedures in teaching practical vocational skills to people with schizophrenia (Kern et al., 2005; Kern, Green, Mintz, & Liberman, 2003; Kern, Liberman, Kopelowicz, Mintz, & Green, 2002). More recently, Fillingham and colleagues have conducted a series of studies examining EL in the rehabilitation of naming abilities in people with post-stroke aphasia (Fillingham, Sage, & Lambon Ralph, 2006; Fillingham, Sage, & Lambon-Ralph, 2005). In this population it appears that although EL is no more effective than control procedures, participants express a clear preference for the EL approach. Rather than providing an exhaustive review of all EL studies, the focus of the next section will be on studies of errorless methods for promoting learning in people with memory disorders, as these have the strongest relevance for applying the technique to PM.

Baddeley and Wilson (1994) examined the impact of EL on verbal learning in 16 people with severe memory impairment. In this study, EL involved providing participants with a two-letter word stem followed by the to-be-learned words and asking them to write them down (i.e. “I’m thinking of a word beginning with BR and that word is “BREAD”. Please write that down”). The control, “Errorful” condition involved giving participants a word stem, asking them to guess the to-be-learned word, correcting that guess, and asking them to write down the target word (i.e. “I’m thinking of a word beginning with BR, please guess what it is” [BREAK? BROOM?] “No it’s BREAD, please write that down”). Every memory-impaired participant learned more words under errorless conditions than errorful conditions, and the “Errorless Learning Advantage” (ELA) was disproportionate to that seen in neurologically healthy young and elderly control

groups. Furthermore, forgetting rates were reduced following EL. Wilson et al. (1994) demonstrated the clinical applications of these findings in a series of case studies. They showed that EL methods were more effective than control methods in rehabilitating naming skills in a person with object agnosia, learning to program an electronic memory aid, learning novel face-name associations, new facts, names of rehabilitation ward staff and items of orientation information.

Subsequent group studies reported in Evans et al. (2000), however, found more mixed results. Benefits for errorless over control learning were identified only for learning names under cued recall conditions or free recall conditions when imagery mnemonics were also used. EL was no different than control for a series of route learning tasks, nor for the task of programming an electronic organiser. The overall conclusion drawn by Evans et al. was that EL would improve performance on any task where implicit memory is used during learning and when the test format facilitates implicit retrieval, however the parameters of these conditions remain to be clearly outlined.

There have been several independent replications of Baddeley and Wilson's original finding (Squires, Hunkin, & Parkin, 1997; Tailby & Haslam, 2003), and a meta-analysis of 8 group studies of EL versus control involving a total of 168 participants reported a large effect size (where $d=.87$, 95% CI: 0.1-1.64; Kessels & de Haan, 2003). More recent studies have also reported beneficial effects of EL on learning virtual reality routes in people with memory impairment (Lloyd, Riley, & Powell, 2009), and in verbal learning in children with brain injury (Haslam, Bazen-Peters, & Wright, 2012). Several studies have also investigated EL approaches in people with progressive memory impairment, including Alzheimer's disease (Bier et al., 2008; Clare et al., 2000; Clare, Wilson, Carter, Hodges, & Adams, 2001; Haslam, Moss, & Hodder, 2010), frontotemporal dementia (Frattali, 2003) and semantic dementia (Jokel & Anderson, 2012). The majority of these studies have found beneficial effects of EL.

Based on the above studies, it seems reasonable to conclude that Errorless Learning is superior to control Errorful Learning procedures in helping people with memory impairment to acquire new skills and learn new information. Several recent studies have moved on to compare EL with other established cognitive rehabilitation techniques, to examine its *relative* efficacy. The findings from these studies are less consistent, and as such will be reviewed in some detail.

Dunn and Clare (2007) sought to examine the relative contributions of varying error rates and effort during learning of face-name associations in a study of ten people with mild dementia. Their four learning conditions were therefore classified as errorless-high effort, errorless low-effort, errorful-high effort and errorful-low effort. No significant differences in performance were observed between any of the conditions on any of three different measures (free recall,

cued recall and recognition). The authors reached the conclusion that error reduction is less important for learning in dementia than other populations, and stated that other means of improving learning such as cognitive effort should be considered. Given there were no benefits from the effort manipulation in this experiment, however, and that there was no formal comparison with other patient groups, this does seem like a rather speculative conclusion. Furthermore, the experiment did not incorporate any control regarding the duration of exposure to correct information, and from the descriptions given, it appears that at least one errorless condition may have been associated with a significantly reduced presentation duration relative to the others.

Haslam, Hodder and Yates (2011) also reported mixed results regarding the efficacy of EL compared with other memory rehabilitation techniques. In one experiment they found that in healthy participants, both EL and Spaced Retrieval (SR) methods were superior to trial-and-error learning for face-name associations. However, performance on EL and SR was equivalent, and there was no benefit of combining EL and SR. Furthermore, in subsequent experiments with people with either ABI, or dementia, only SR was associated with improved recall. These latter results are clearly contrary to those from many other studies, and the authors acknowledge that their particular EL protocol was rather passive in comparison with the other conditions, which may have contributed to this surprising finding. Furthermore, the SR condition was advantaged relative to the other conditions in having a shorter delay between the final presentation of correct information in the study phase and the subsequent test phases (5 minutes in SR as opposed to 8.5 minutes in all other conditions). The authors note that this was unlikely to have large impact on the results, and the study was commendable for having carefully balanced the time from the *start* of the study procedure and test phase between conditions. However, given SR was the only condition associated with improved performance in the clinical groups, and that it fundamentally differed from all other comparison conditions, this reasoning seems rather unfounded. A further unacknowledged confound is that the trial-and-error and SR conditions involved presentation of face stimuli for much longer than the EL condition (13sec vs 3sec each), meaning that when these appeared at test, they would be much more familiar than the errorlessly learned faces. Though it is hard to predict the impact of this bias on the pattern of results (e.g. would the relatively less familiar errorlessly learned provide a lesser trigger for the associated names, would there be a greater proportion of “don’t know” responses?”), it does contribute to an overall impression that the implementation of EL and SR in this study was somewhat flawed.

On the surface, the comparison studies appear to imply that EL methods are not as beneficial as the early reports indicated⁸. Indeed, a recent review by Middleton and Schwartz (2012) sounded a note of caution to this effect. However, a critical analysis of the methodologies employed in the comparison studies makes it very clear that their results are far from conclusive. Future studies aimed at comparing learning methods would benefit from the inclusion of more careful control conditions for example by comparing “errorless spaced retrieval” with “errorful spaced retrieval”, as well as “errorless non-spaced retrieval” and “errorful non-spaced retrieval” conditions, and reporting the error rates, exposure duration, and study-test delay for each.

One of the many interesting points made by Middleton and Schwarz in their review is that procedures designed to minimise errors may be associated with unintentional reduction in opportunities for retrieval practice⁹, and this means that in some circumstances, EL procedures may be detrimental to learning. This is a very apposite statement, and one that harks back to the clinical guidance offered by Wilson et al. (1994), that the EL principle should be incorporated into an active learning paradigm, as this would likely involve maximising opportunities for successful retrieval. Despite these cautions, there seem to be strong empirical grounds to state that *all other factors being equal*, procedures that minimise errors during learning are likely to improve subsequent memory performance.

2.7.4. Mechanisms of the Errorless Learning Advantage

Wilson and Baddeley’s rationale for applying EL to memory rehabilitation was simple: in trial-and-error learning, one needs to learn not only the correct response, but also that one’s errors are indeed errors. By devising a training procedure that eliminates errors as far as possible, the memory demands of the task are reduced, and performance is improved. They also refer to the Hebbian learning maxim (Hebb, 1949) that “cells that fire together wire together”. When erroneous responses occur, they are associated with the spatiotemporal context in which they were produced, and hence they are more likely to be reproduced in similar future contexts. Importantly there is no mechanism for identifying and removing errors (McClelland, Thomas, McCandliss, & Fiez, 1999).

A number of studies have addressed the question of whether the Errorless Learning advantage stems from implicit or explicit memory processes. As people with amnesia have severely

⁸ There are even some areas in which EL has been found to result in significantly worse performance relative to control approaches, for example, in source memory tasks (Cyr & Anderson, 2012). However this study concerned neurologically healthy adults in whom EL effects are predicted to be smaller than in people with memory impairment.

⁹ Retrieval practice being an established principle associated with improved performance (Roediger & Payne, 1982).

impaired explicit memory, Baddeley and Wilson (1994) thought that the ELA was mediated by implicit memory processes. However, other researchers argued that the ELA resulted solely from residual explicit memory capacities. Hunkin, Squires, Parkin and Tidy (1998) taught word lists to people with moderate-severe memory impairment under Errorless and Errorful Learning protocols, before testing their memory in two formats, first fragment completion and then stem completion. If for example the learned word was ARTIST, the fragment completion task would test memory for the previously learned word with the array “_ _ T _ S _”, which was considered an indirect test, and to tap implicit memory. The subsequent stem completion test would present the array “A R _ _ _”, and this was considered a more direct and explicit cued recall test¹⁰. Hunkin et al. found an ELA only in the stem completion task, and interpreted this to mean that the ELA relies upon residual explicit memory. However, Page et al. (2006) present a strong argument to the contrary. They stated that the implicit task used by Hunkin et al. (1998) was designed such that it was insensitive to implicit memory for prior errors. Say for example that the errors ‘ARCHES’ and ‘AROUND’ had been made while learning the word ‘ARTIST’. The stem “A R _ _ _” primes both the correct response and the errors, whereas the relevant fragment “_ _ T _ S _” primes *only* the correct response. No advantage based on avoiding implicit memory of errors could therefore be detected. Page et al. (2006) also found that people with moderate memory impairment and hence some remaining explicit memory, showed no greater advantage from Errorless Learning than people with very severe memory impairment who had little to no residual explicit memory. This again supports the hypothesis that implicit memory is sufficient to produce the ELA.

Other studies have investigated the mechanisms of EL in normal ageing. Anderson and Craik (2006) compared errorless versus control Errorful Learning conditions in young and older adults using a process dissociation procedure that allowed for measurement of recollection (explicit memory) and familiarity (implicit memory). They found that EL was associated with reduced familiarity in both age-groups, but reduced recollection in younger adults only. On the basis of this result, Anderson and Craik stated that EL is not helpful for people with intact explicit memory because the advantage of having reduced interference from prior errors is outweighed by the disadvantage of having an unelaborated learning process in EL relative to Errorful Learning. In older adults however, who have poorer recollection, EL is beneficial, as it decreases the negative impact that prior errors have on familiarity. This is therefore consistent with the ELA stemming from implicit memory.

¹⁰ Note that both stem and fragment-completion tasks have been considered measures of implicit memory elsewhere (e.g. Squire, Shimamura, & Graf, 1987). Hunkin et al. (1998) acknowledge this point, but consider for the current purposes that stem completion is a relatively *more* explicit test than fragment completion. (Squire & Zola, 1996)

Consistent with these results, Anderson et al. (2012) studied healthy older adults classified as high/low in medial temporal lobe (MTL) and Prefrontal Cortex (PFC) functioning according to performance on standardised neuropsychological tests, and found an ELA in tests of free and cued recall, but not recognition. Furthermore, the ELA in cued recall was marginally associated with MTL functioning, with those with worse memory showing a larger ELA. This is consistent with there being a mnemonic basis to the ELA, though it does not specifically address the implicit/explicit distinction.

There have also been a small number of neuroimaging studies in neurologically healthy volunteers that have attempted to address questions about the neural basis of the ELA. Event-related potential (ERP) studies have identified modulation of right frontal ERPs, and increased error-related negativity during retrieval of words learned under errorless conditions compared with trial-and-error conditions. In a subsequent fMRI study, Hammer, Templemann and Münte (2013) identified increased frontoparietal activations associated with control (errorful) learning compared with EL. Furthermore, Hammer et al. (2011) conducted a transcranial direct current stimulation (tDCS)¹¹ study where cathodal, anodal and sham stimulation were applied to left dorsolateral prefrontal cortex. This area was selected on the basis of its involvement in encoding particularly of verbal material (c.f. the Hemispheric Encoding/Retrieval Asymmetry model; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994). They found that cathodal stimulation impaired performance after Errorful Learning only – there was no impact of tDCS on performance after EL. This implies that the PFC has a role in learning with errors that is either absent or substantially smaller in EL.

The conclusions from neuropsychological studies of the ELA are therefore that people with compromised explicit memory are more affected by errors made during the learning process, and that implicit memory is sufficient to produce an ELA. The results of the neuroimaging and neurostimulation studies suggest that Errorful Learning places greater demands on frontal/executive resources than Errorless Learning, and it seems likely that these executive demands are necessary to identify and reject prior errors. These two sets of conclusions are not entirely incompatible; the detection of frontal lobe involvement in Errorful Learning does not exclude the possibility that the *ELA* results from implicit memory; nor does the observation of a larger ELA in people with poorer MTL function necessarily mean that the ELA is *not* mediated by frontal (or other) substrates. However, there are some inconsistencies in the predictions that follow from these observations. For example, on the basis of the neuropsychological findings one might predict that neuroimaging studies would have found

¹¹ tDCS is a non-invasive technique that briefly alters neural firing rates in relatively discrete brain areas by passing a current through electrodes in contact with the scalp. Cathodal stimulation causes a decrease in firing rate, and anodal an increase.

increased MTL activation patterns for Errorful Learning conditions, reflecting their higher mnemonic demands, but this was not so. Equally, given the neuroimaging findings, one would expect that people with weak executive functioning or frontal lobe damage would show a larger ELA than those with better executive functioning/no FL damage, since the Errorless condition reduces executive-mnemonic demands of the task that would be particularly challenging for the former group. However, Anderson et al.'s (2012) low-frontal lobe groups did not show a larger ELA than the high-frontal groups. Furthermore, Baddeley and Wilson (1994) divided their amnesic patients into groups according to their executive abilities (no executive dysfunction, borderline, and dysexecutive) and found an ELA of a similar magnitude in each group. It is certainly not unusual for there to be discrepancies between neuropsychological and neuroimaging studies (Shallice, 2003), and it is of course possible that there are several mechanisms underpinning the ELA, and/or that they might differ between patients and controls. Clearly there are issues for further research to address, and imaging/stimulation studies within clinical groups may serve to illuminate some of these issues.

2.7.5. Applying Errorless Learning to Prospective Memory tasks

It stands to reason that an intention for future action that is clearly stored in memory is more likely to be acted upon than an intention that has been incorrectly stored or weakly remembered. Hence, EL could improve PM performance relative to Errorful learning by decreasing the likelihood of intentions being forgotten “outright”, before the appropriate retrieval opportunity is presented. This would be a relatively uncontroversial finding, albeit one with applications for clinical work with people with memory disorders. A more exciting prediction could be made, however, on the basis of the ELA research previously discussed. Specifically, if implicit processes are sufficient to produce an ELA, and/or if EL is successful due to its decreased mnemonic demands, one might predict that errorlessly-learned intentions would be more amenable to timely retrieval than intentions formed with errors during learning – even if the intention was adequately retained in both encoding conditions. Given the previously described differential reliance of Time and Event-based PM tasks on executive and mnemonic processes respectively, one might further predict *that mnemonic strategies such as EL would be more effective for Event-based PM performance than Time-based PM performance.*

If the ELA were frontally mediated, however, as suggested by the tDCS and fMRI studies, then the predictions would be less secure. One might predict that as errorlessly-learned intentions necessitate reduced recruitment of frontal/executive resources during retrieval, then such resources would be “freed up” and hence performance on the Time-based task would benefit

over and above the Event-based one. However, as stated previously, these frontal-ELA results may be specific to neurologically healthy participant groups, and it is certainly more sensible to base the predictions on the more established results from groups of people with memory disorders¹².

Furthermore there is no published research on the effect of EL on executive task performance. On a theoretical basis, a benefit may be expected from EL on executive tasks if task performance were supported at least in part by implicit memory processes (e.g. if the task involved learning motor sequences that would have a procedural component), but this seems unlikely to apply to Time-based PM tasks, due to their reliance on more active monitoring.

2.8. Aims of the Study

A set of hypotheses was generated on the basis of the literature reviewed. As Errorless Learning has been found to result in improved memory for verbal material in comparison with Errorful Learning particularly on tests of cued recall, there should also be an ELA for PM retrieval, under certain conditions. Specifically, an ELA would be expected on those PM tasks that have been shown to rely upon mnemonic processes. As Event-based PM tasks have greater mnemonic and lesser executive demands than Time-based PM tasks, Hypothesis A was that *there will be a significantly greater ELA for Event-based PM tasks in comparison with Time-based PM tasks*.

As the evidence suggests that the ELA stems from implicit memory (or avoiding the need to recruit further explicit memory to screen out errors), rather than any controlled search process, and as any additional search processes would seem likely to have a detrimental impact on Ongoing Task performance (i.e. the background task within a PM paradigm), Hypothesis B was that *any ELA will occur without any associated detrimental impact on Ongoing Task performance*.

Little is known about metacognitive aspects of PM performance, particularly in people with Memory Disorders. However, the study of such metacognitive awareness could potentially be theoretically and clinically informative. Consequently, the study included measures of awareness of different aspects of the PM task.

This study investigated these hypotheses within a 2x2 factorial experiment employing a within-subjects design, with counterbalancing to avoid confounding order effects. A group of people with memory impairment were taught the instructions for four PM tasks (two Time-based, two Event-based) under two different "Encoding" conditions. The instructions for one EBPM and one TBPM task were learned in a way that minimised the occurrence of errors during learning (Errorless Learning), and the remaining tasks in a way that ensured the production of errors

¹² Note also that these studies were published after the project was designed.

(Errorful Learning). After each encoding procedure, participants completed a task that comprised an attentionally demanding sentence verification “Ongoing Task”, with one of the embedded Prospective Memory tasks. This process was then repeated until all conditions had been completed. Statistical analyses then examined the effects of encoding and task type on PM task performance, Ongoing Task performance, and awareness of PM performance.

3. Method

3.1. Ethical Review

The study was reviewed and approved by an NHS Research Ethics Service Committee (London – Camden and Islington branch, REF 12/LO/0310), the Psychological Medicine Clinical Academic Group within King's Health Partners, and South London and Maudsley NHS Foundation Trust (SLaM) Research and Development Office. All participants gave written informed consent (see Appendix, Section 7.1. for procedures regarding mental capacity to consent).

3.2. Participants

3.2.1. Recruitment

Participants were recruited from the Neuropsychiatry and Memory Disorders Service at St Thomas' Hospital in South London. The clinic accepts referrals for general neuropsychiatric services (i.e. for the psychiatric and/or behavioural aspects of medical conditions, especially those with neurological involvement), and has also has a memory disorders service which provides specialist assessment and treatment of conditions associated with memory or other cognitive impairment, including from brain injury and dementia. To identify patients eligible for the study, the researcher attended and screened records for clinics held by two Consultant Neuropsychiatrists over a period of nine months. No data regarding ineligible participants were recorded, so the total number of records screened cannot be calculated. However, an estimate can be obtained by calculating the number of appointments per clinic (approximately 7) multiplied by the number of clinics attended (47) divided by the likely number of repeat appointments a patient was likely to have during that time (2-3). This suggests that 110-164 patients were screened for eligibility. Seventeen were considered eligible according to the criteria specified below.

3.2.2. Eligibility Criteria

Patients meeting the following criteria were considered eligible for participation in the study:

- Aged 18-70 years
- Sufficient fluency with written and spoken English to complete the experimental task
- Presence of memory impairment, defined as performance on verbal memory tests at least 1.5 standard deviations lower than would be predicted on the basis of IQ score (see below for details).
- Evidence of a neurological basis for the memory impairment.
- A minimum period of 12 months elapsed from the onset of the memory problem.

- Memory problem not of a progressive nature
- No documented history of Learning Disability, nor current verbal IQ < 70, to ensure ability to complete the experimental task.

Seventeen patients were deemed eligible and three declined to participate. In one case this was due to stressful additional health problems. The other two cases initially agreed but later declined or could not be contacted. This left a final sample of fourteen.

3.2.3. Sample Characteristics

3.2.3.1. Demographics

The sample comprised fourteen people (twelve males, two females), with a mean age of 53.93 years (SD 8.27, range 38-69). The group had an average of 11.86 years of education (SD 2.19, range 9-17). Four participants had received a primary education, two had obtained qualifications of secondary education, four had obtained A-levels or equivalent, three had degree-level education, and one had a postgraduate degree. Occupational classifications from the Office for National Statistics (ONS) categorised two participants as previously engaged in higher managerial or professional positions, six as intermediate occupations (e.g. clerical or service roles), and six as in routine or manual occupations. Eight participants were currently unemployed and in receipt of disability benefits, two had retired from work on medical grounds, two had retired prior to the onset of their memory difficulties, and two were in part-time employment. According to ONS ethnicity classifications, twelve participants were White British, one Black African, and one Black Caribbean.

3.2.3.2. Aetiological Information

The aetiology of patients' memory difficulties were in six cases cerebrovascular disease, three cases cerebral hypoxia, two cases temporal lobe epilepsy, and one head injury and small vessel disease. In two cases the aetiology was uncertain, with one case being either a stroke or hypoxia secondary to status epilepticus, and another neurological damage in the context of long-term poorly controlled diabetes (for more detail see Appendix section 7.2.).

3.2.3.3. Definition of Memory Impairment

Memory impairment was defined as a discrepancy of 1.5 standard deviations between scores on standard tests of intellectual ability and memory administered routinely within the clinic. The precise tests used varied slightly between patients, for example where patients were not native English speakers or where a certain test had been administered several times, but there were always results from recall and recognition tests of visual and verbal memory. The decision was made to use existing test scores, rather than to re-administer tests for the

purposes of the project because recent test results were known to be available for the vast majority of clinic patients, and it reduced the burden of participation for those volunteering.

There are several potential methods of establishing current intellectual ability for these purposes, each with their limitations. For example, the most widely used way of estimating premorbid IQ is by measuring reading of words with irregular pronunciations, as this skill is both highly correlated with intellectual ability in the general population, and largely resistant to decline in neurological conditions. This method is, however, inappropriate for people with dyslexia or for non-native English speakers. Measures of current, particularly non-verbal, intellectual ability may be preferable in such cases. These, however, can also be affected in neurological conditions. In particular, fluid intelligence is frequently affected by brain injury particularly to the frontal lobes. As the current sample included people with a history of dyslexia, and non-native speakers of English, and people with additional frontal executive difficulties, the pragmatic decision was made to use different measures of current ability on a case-by-case basis. The estimated Full-Scale IQ score (FSIQ) from the National Adult Reading Test (Nelson & Willison, 1991) was used in three cases, the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001) in one case, and from the Test of Premorbid Functioning (TOPF; Wechsler, 2011a) in two cases. The 2-subtest FSIQ score from one of two versions of the Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler, 2011b; WASI; Wechsler, 1999) was used in four cases, and the Matrix Reasoning t-score alone in four cases. It is clearly not ideal to measure intellectual ability on the basis of one WASI/WASI-II subtest only, however, this measure is the subtest does have a strong correlation with WASI FSIQ ($r=.84$), and is relatively less influenced by cultural and educational factors than the verbal subtests.

Memory performance was measured with the overall memory score from the Doors and People battery (Baddeley, Nimmo-Smith, & Emslie, 1994) in twelve cases, the composite verbal score in one case, and a composite from the story recall and figure recall subtests from the BIRT Memory and Information Processing Battery (BMIPB; Coughlan, Oddy, & Crawford, 2007) in one case¹³.

The inclusion criteria of memory performance at least 1.5 standard deviations below intellectual ability, indicating impairment relative to expected premorbid ability, was therefore met ($M z = -2.55$, $SD .69$, max. -1.67 , min. -4). The group's intellectual functioning was consistent with the population mean ($M z = .27$, $SD = .81$, max. 1.67 , min. $-.75$), and their memory functioning was impaired ($M z = -2.26$, $SD = .69$, max -3 , min -1.33). Note that two participants' memory scores would be classified as in the borderline range (9th centile).

¹³ These deviations were due to differences in the test batteries in one case, and due to sparing of visual recognition memory in one case.

As would be expected in a mixed-aetiology group such as this, participants' impairments were not restricted to the domain of memory. Indeed, five participants had borderline scores on at least one test of executive function. Further detail on the sample is presented in Section 7.2, Appendix.

3.3. Study Design and Counterbalancing

The study was a 2 (encoding condition: Errorless Learning, Errorful Learning) x 2 (PM task type: Time-based, Event-based) factorial within-subjects experiment. All participants therefore took part in four experimental conditions: Errorless Learning of an Event-based PM task, Errorless Learning of a Time-based PM task, Errorful Learning of an Event-based PM task, and Errorful Learning of a Time-based PM task. To minimise practice effects, parallel versions of the PM tasks were developed (PM task versions A, B, C, D; A and C being Event-based, and B and D being Time-based).

For between-group studies, even if good group matching is achieved, individual difference factors mean that the between group variance is inevitably far higher than within group variance, necessitating substantially larger samples sizes, which was not possible for this study. Hence the within-subjects design was preferred.

Full counterbalancing of all experimental factors may have lead to some contamination of the encoding conditions. Specifically, switching between the instructions to "only respond if you're sure you're right" versus "if you're not sure, have a guess" within one session may have made it difficult to control the presence or absence of errors in the later part of the session. Therefore, learning condition was counterbalanced on a between-session basis, so that half of the participants underwent Errorless Learning first, and half underwent control learning first. PM task type was crossed to control for order effects, with the restriction that only one Time-based and one Event-based task occurred in each session. This resulted in 16 unique administration sequences (see Table 1.2.), 14 of which were used in the study.

Table 1.2. Task administration order. Note that grey/white shading refers to learning condition, whereby grey shading indicates Errorless Learning and white control.

Participant	SESSION 1		SESSION 2	
	Task 1	Task 2	Task 3	Task 4
1	Event-based A	Time-based B	Event-based C	Time-based D
2	Event-based A	Time-based D	Event-based C	Time-based B
3	Event-based C	Time-based B	Event-based A	Time-based D
4	Event-based C	Time-based D	Event-based A	Time-based B
5	Time-based B	Event-based A	Time-based D	Event-based C
6	Time-based B	Event-based C	Time-based D	Event-based A
7	Time-based D	Event-based A	Time-based B	Event-based C
8	Time-based D	Event-based C	Time-based B	Event-based A
9	Event-based A	Time-based D	Time-based B	Event-based C
10	Event-based A	Time-based B	Time-based D	Event-based C
11	Event-based C	Time-based D	Time-based B	Event-based A
12	Event-based C	Time-based B	Time-based D	Event-based A
13	Time-based D	Event-based C	Event-based A	Time-based B
14	Time-based D	Event-based A	Event-based C	Time-based B

3.4. Procedure

3.4.1. Overview

All participants gave informed consent before the experimental tasks began. They were informed that the general aim of the study was to investigate the effects of different learning methods on memory, but any information regarding the specific experimental hypotheses was withheld, to minimise any bias that could result from such knowledge.

Participants completed each of the four experimental tasks in either one or two sessions. The original intention had been to separate each session with an interval of one week; however, a number of participants requested one longer session rather than two shorter ones (typically these participants who would need to travel a significant distance to the clinic). As the proportion of eligible participants from patients attending the clinic was rather low, and the likelihood of incurring missing data for the second session was increased by holding the

sessions on different days, such requests were accommodated. This meant that six participants completed the experimental tasks in one day, and eight across two days, one week apart.

Each participant completed the four conditions in one of the orders specified in Table 1.2. Each condition included four distinct phases:

- 1) **Encoding:** Presentation of instructions with vanishing cues (VC), in 3-5 cycles separated by a distracter digit span task, until the criterion of correct recall after digit span task was reached.
- 2) **Delay:** completion of questionnaires for set 4-minute period¹⁴
- 3) **PM task:** duration 16.5 minutes
- 4) **Post-test questions:** these assessed memory for the PM and ongoing task instructions, along with questions aiming to obtain a self-estimate of performance.

At the end of the study, participants were given a brief verbal summary of the study's aims, thanked for their time and support of research within SLAM, and reimbursed for their travel costs at a flat rate of £10 per session. Sessions for all but one participant took place at the Neuropsychiatry and Memory Disorders Service clinic base at St Thomas' Hospital. The remaining participant preferred to be visited at home.

3.4.2. Encoding Procedures

Both Errorless and Errorful Encoding conditions involved participants learning the instructions for one of the PM tasks as a single sentence, e.g. "Press the red (blue) key when you see the word "tigers" ("hammers")", or "Press the blue (red) key every other minute, starting at 1:00 (2:00)".

Initially, the experimenter stated the task instructions for the ongoing task, and demonstrated it for approximately 5 trials. The participant then practised for a similar duration. This process was then repeated for the PM task, but without making specific reference to the target word or time interval (e.g. "every now and then you'll have to do a different task, and for that task you press this button. We'll go through the details in a moment").

The PM task instruction was then learned using one of two Vanishing Cues procedures implemented using Microsoft PowerPoint along with prompts from the experimenter (see Figure 1.2. for a summary).

¹⁴ Note that this delay is somewhat shorter than that used with healthy subjects in the experimental literature, which are often 10-15 minutes, but it is consistent with the few studies of PM in people with memory impairment.

3.4.2.1. Errorless Learning

The PM instruction sentence was displayed on a laptop monitor in white point 20 Arial font lettering on a black background. The participant was asked to read it aloud. The instruction sentence was then displayed with the final word deleted, it being represented only by blank underlining approximate to the missing word's length (see Figure 1.2). The participant was again asked to read it aloud, and to fill in the gaps, but *only if they were sure they know the answer*. They were instructed that if they were unsure of the response, they would be provided with a clue leading to the correct answer. To facilitate an active approach within the task, clues for certain words were in the form of descriptions (e.g. for the word "tigers", the description was "they're wild animals, a type of "big cat" that has stripes, the name has six letters and it starts with T". Such descriptions were also given for the following content: hammers, red, blue, 1:00, 2:00). This procedure continued, removing one word at a time, until the sentence could be recited in response to a series of blank lines representing each word. A brief distracter task was then completed, specifically 10 trials of a digit span task. This served to prevent continuous rehearsal of the task instruction. The Vanishing Cues procedure then resumed, with the first 5 words of the instruction displayed initially for the participant to complete, before words were again progressively removed. The digit span task was then repeated, and the Vanishing Cues procedure resumed, beginning with only the placeholders remaining (see Fig 1.2). If at any point the participant was not confident in producing a correct response, a greater proportion of the sentence was displayed, before the process of progressive fading began again.

3.4.2.2. Errorful Learning

The Errorful Learning procedure mirrored the Errorless Learning procedure, apart from the following characteristics. Firstly, a "forced error" was elicited by asking participants to guess the target word/number. A second error was introduced by changing the stated colour of the key press response between the first and second repetitions. Finally, rather than discouraging guessing, guessing was encouraged within the instructions. Aside from these manipulations, the previously outlined Vanishing Cues procedure was followed. If at any point a spontaneous error was made, the procedure reverted to the previous screen and the whole instruction was repeated. This was to ensure that exposure to the correct information was equivalent between the encoding conditions.

The criteria for completing the encoding phase was accurate recall of the instruction in response to the blank sentence placeholders only, after a minimum of three VC cycles. The number of trials needed to meet the learning criterion was recorded, along with any errors made. See Figure 1.2 for a flowchart detailing the Errorless and Errorful encoding procedures.

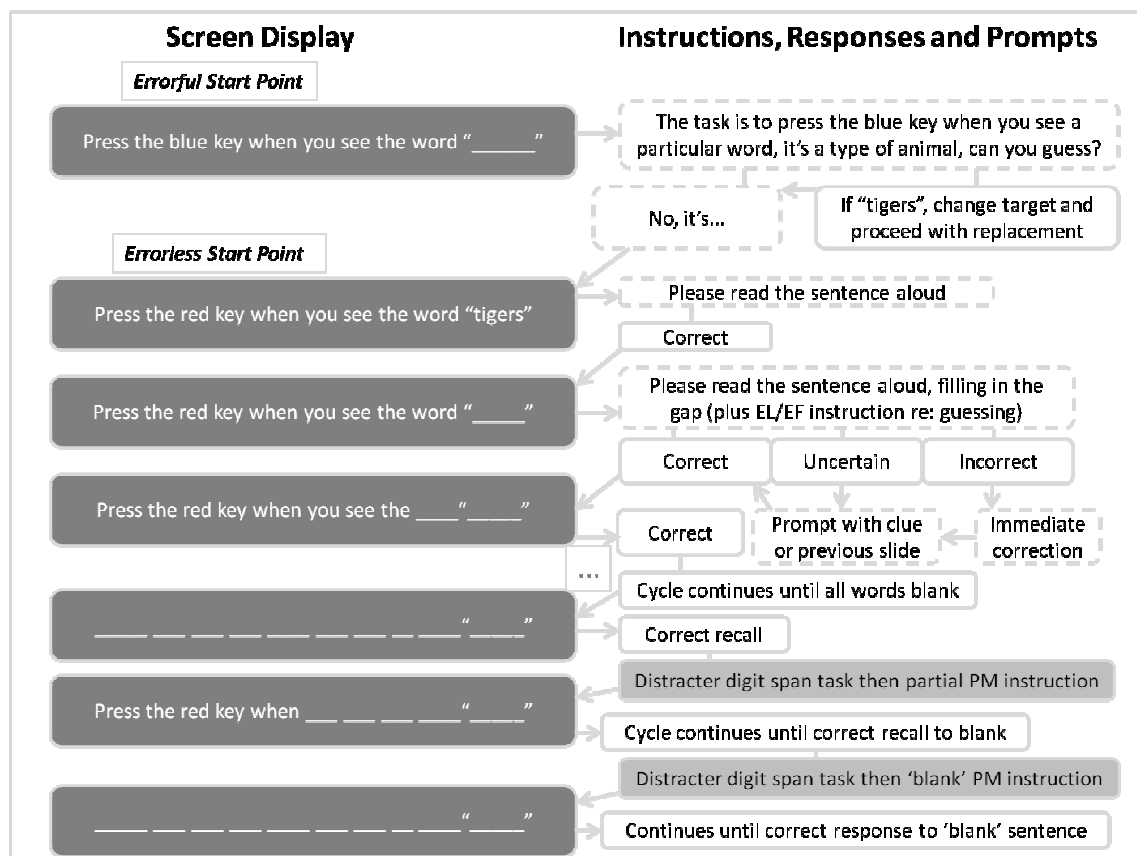


Figure 1.2. Flow chart detailing procedures for Errorless and Errorful Encoding

3.5. Measures

3.5.1. Prospective Memory Paradigm

Consistent with the experimental tasks described in the introduction, the PM task involved an attentionally demanding Ongoing Task with a further PM task to be performed on an infrequent basis. The task ran on a Dell Latitude D520 laptop computer. Throughout, participants responded using one of four clearly labelled keys located towards the lower right of the laptop keyboard. The leftmost was marked with a clock-face symbol and when pressed, this displayed the time elapsed from the start of the task. Adjacent to this was the PM response key, marked with either a red or blue sticker depending on the task version. To the right of this, there were two keys marked "T" and "F", with which true/false responses to the ongoing task were made. See Figure 1.3. for an illustration of the task.

3.5.1.1. Ongoing Task

The ongoing task was the same for all four versions of the paradigm. It was based upon the Speed of Comprehension test from the Speed and Capacity of Language Processing Test

(SCOLP; Baddeley, Emslie, & Nimmo-Smith, 1992), also known as the “silly sentences test”¹⁵. Participants were presented with a series of sentences which they are asked to read and then make a decision regarding whether or not the sentence is true (e.g. “apples are fruit”, “desks can be bought in shops”) or false (e.g. “beef steaks are fruit”, “physicists can be bought in shops”). There were 404 such sentences available for use in the task, an equal proportion being true and false in content, and these were selected at random for the 16.5-minute duration of the task. Sentences were displayed in a white sans serif font approximately 10mm in height, on a black background. The task was self-paced, with participants being instructed to respond as quickly as possible, whilst avoiding errors. The outcome measures were the number of sentences rated, the percentage correct responses, and median response times for both correct responses and errors.

3.5.1.2. PM tasks

There were four versions of the PM task; two being Event-based, and two Time-based. To facilitate comparisons between the tasks, they were constructed such that each task lasted 16.5 minutes and that perfect performance necessitated 8 responses in each condition. The instruction sentences were balanced so that each contained ten words, and instructions for clock-monitoring were included in all conditions even though this was a task-relevant activity in the Time-based tasks only.

The Event-based task was to press the designated key when a designated word appeared within a sentence (Either: Press the red key when you see the word “tigers”, or Press the blue key when you see the word “hammers”). The target word appeared at pseudorandom intervals, the order of which was fixed for all participants (PM trials in Task A occurred on the trial subsequent to each of the following time-points: 1:25, 2:17, 4:12, 6:09, 7:38, 9:12, 11:23, and 12:07; and in Task C at 1:55, 3:46, 7:43, 8:16, 9:25, 11:14, 12:36, and 13:30). The outcome measures in this task were correct PM responses and erroneous PM responses. Though response times were recorded, due to the low number of trials, they were not analysed further.

The Time-based PM tasks were to press the designated key every other minute, starting at 1:00 (version B), or to press the designated key every other minute starting at 2:00 (version D). This meant that as with the Event-based task, optimal performance would mean eight PM responses were made in each condition. Time-checking behaviour served a more obvious purpose in this Time-based condition, which is to assist in accurate timing of the PM response.

¹⁵ This name refers to the occasionally comical nature of the false sentences in particular, which were created by mixing up the beginning and end of true sentences.

The timing and frequency of time-check responses was recorded, but not used in the analysis. PM responses within a window of 90 seconds of a target were scored as correct.

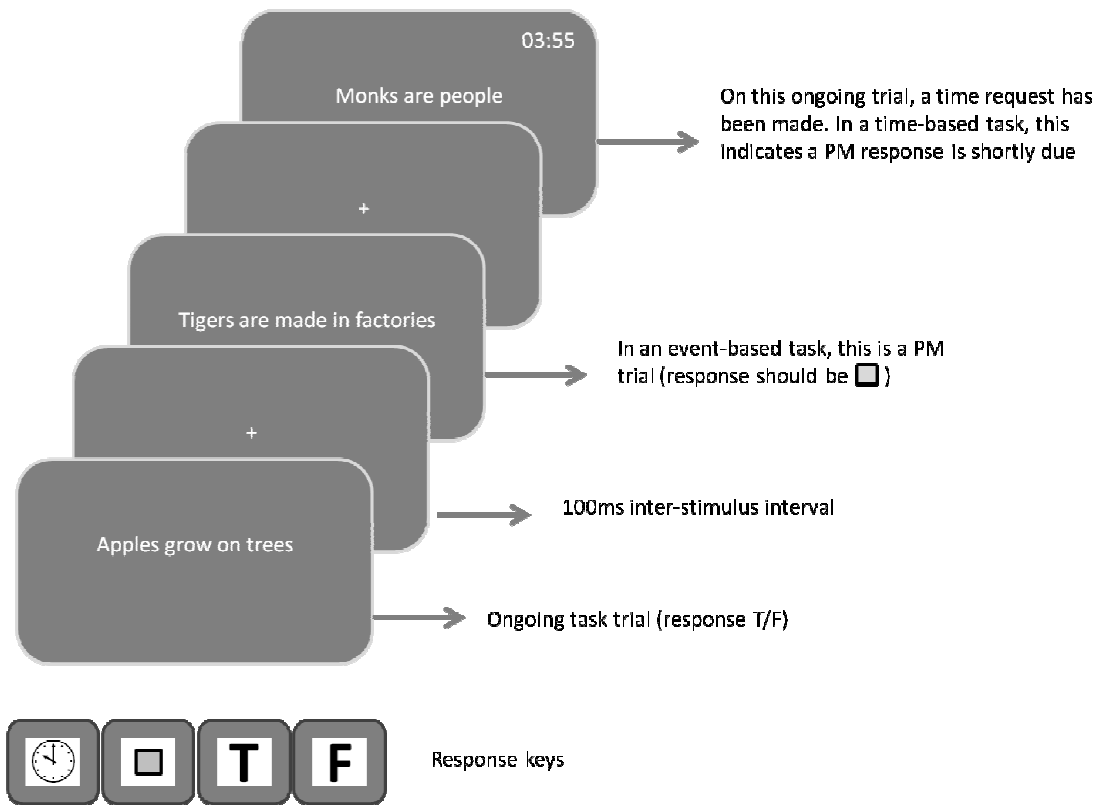


Figure 1.3. Illustration of the trial sequence and response format of the PM task.

Note that examples of both Event-based and Time-based tasks are included, though in reality these would not occur within the same sequence.

3.5.1.3. Post-Test Awareness Questions

At the end of each task participants were asked to recall the instructions for both the ongoing and PM tasks. In the event of erroneous or “don’t know” responses, cued recall was tested (i.e. you were to do something when you saw a particular word, what was the word/task?)

After the Event-based task, participants were asked to estimate how many times the target word appeared, and how many times they pressed the designated key in response to it. In the Time-based tasks, they were asked how many times they pressed the designated key, and how many times they *should* have pressed it. This allowed for creation of four variables relevant to self-perception of performance: target awareness (i.e. the response to the question asking “how many times should you have pressed red [or blue]?” divided by the number of targets presented), response awareness (i.e. response to question “how many times did you press red/blue?” divided by actual number of PM responses), subjective accuracy, expressed as a

proportion (i.e. number of times pressed red divided by number of times should have pressed red), and discrepancy (objective accuracy minus subjective accuracy).

3.5.2. Background Neuropsychological Measures

Scores on standard tests of cognitive functioning were obtained, with participants' consent, from clinical records. In all cases this included tests of recall and recognition of verbal and nonverbal material (e.g. from the Doors & People Test; Baddeley, Nimmo-Smith & Emslie, 1994) or a similar measure, along with a measure of current or estimated premorbid IQ (see section 2.2.3.3). Additional measures of naming/semantic memory (Graded Naming Test; McKenna & Warrington, 1983) were available for 10 participants, and executive functioning (letter and category fluency, Hayling & Brixton tests; Shallice & Burgess, 1997) were available for all but one.

3.5.3. Questionnaires

To obtain estimates of participants' self-reported cognitive functioning, mood and anxiety, and general experience of problems frequently experienced by people after brain injury participants completed three questionnaires, described below. The results were not used in the analyses but are presented in the Appendix in Section 7.2, to assist in the characterisation of the sample.

3.5.3.1. Prospective and Retrospective Memory Questionnaire (G. Smith et al., 2000).

This 16-item questionnaire includes 8 items focussed on Prospective Memory and 8 items focussed on Retrospective Memory. Each subscale includes four short self-cued and four environmentally cued tasks, each with 2 items referring to short-term and 2 to long-term tasks. Factor analysis has identified that although all items load on a general memory factor, there are additional orthogonal prospective and retrospective factors (Crawford, Smith, Maylor, Della Sala, & Logie, 2003). The reliabilities for total and subscale scores are $>.80$.

3.5.3.2. Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983).

This widely-used 14-item measure includes 7 items that correspond to an Anxiety subscale, and 7 to a Depression subscale. Several studies have suggested this is a valid measure for individuals with brain injury (Dawkins, Cloherty, Gracey, & Evans, 2006; Schönberger & Ponsford, 2010). Crawford et al. (2001) found both subscales and the total score to be internally consistent (Cronbach's alpha .82, .77, and .86).

3.5.3.3. The European Brain Injury Questionnaire (Teasdale et al., 1997).

This 63-item questionnaire lists a range of symptoms that are sometimes experienced by people with brain injury, and requires the respondent to state whether they have experienced that symptom ‘not at all’, ‘a little’, or ‘a lot’ in the previous month. The EBIQ was originally constructed with eight subscales: somatic, cognitive, motivation, impulsivity, depression, isolation, physical and communication, along with a “core symptom” subscale. The measure has adequate internal consistency (median Cronbach’s alpha of .63 for patient self-report) and test-retest reliability (median $r=.76$; Sopena, Dewar, Nannery, Teasdale, & Wilson, 2007). Scores for the most relevant subscales (cognitive and core symptoms) are presented in the Appendix, section 7.2.

3.6. Power Calculation

Power calculations for the primary analyses of the main effect of encoding method on Event-based PM performance, and the interaction between encoding and PM task type, were conducted using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). The calculations were guided by the effect size reported by Kinsella et al. (2007), which is the closest existing study to the one proposed here.

In Kinsella et al.’s study the size of the effect of spaced retrieval on PM performance was $d = 1.25$, which exceeds Cohen’s (1992) threshold of $>.8$ for a ‘large’ effect size. As the proposed study uses a different encoding technique, a more conservative effect size of .8 was carried forward for use in the calculation. On this basis, a sample of 12 participants would give 83% power to detect an effect significant at the 5% level in a one-tailed paired t-test. There was no existing clinical study comparing encoding procedures for different PM tasks on which to base a power analysis for the interaction effect. However, taking a more cautious estimate of a medium effect size, a sample size of 28 would give 81% power to detect a significant interaction within a 2x2 repeated measures ANOVA, presuming a correlation of .4 between the repeated measures. The intention was therefore to recruit and test up to 28 participants, though it was acknowledged that this was an optimistic target. The final sample of 14 gave 88% power to detect large-sized effects, but only 46% power to detect the presumed medium-sized interaction. Therefore, the statistical analyses were restricted to the two paired t-tests that examined the effects of interest, rather than the full factorial ANOVA.

3.7. Planned Statistical Analyses

The planned statistical analyses aimed to: (a) examine the effectiveness of the experimental manipulation in prompting and minimising errors as appropriate, (b) ensure that exposure to the correct information was equivalent between the encoding conditions, (c) to examine the

impact of errors during learning on subsequent PM accuracy for both the Event-based and Time-based tasks, (d) to identify any impact of Encoding Condition or Task Type on Ongoing task performance, and (e) to investigate whether there the experimental manipulations had any impact on participants' awareness of task-related information.

The approach was to initially inspect the distributions of raw data for all dependent variables of interest, by means of boxplots, Q-Q plots, and variance estimates. Where the distributions were approximately normal, parametric statistical tests were used, otherwise, non-parametric equivalents were employed. The differences between conditions were then examined using Analysis of Variance, t-tests, or their parametric equivalents. Effect sizes were also computed.

4. Results

4.1. Preliminary Analyses Concerning the Efficacy of the Experimental Manipulations

This section presents results from a series of analyses conducted to determine the effectiveness of the experimental procedures in manipulating error rates and teaching PM task content, which are both essential for the interpretation of the subsequent experimental results. Analyses are presented for error rates and exposure to correct information during learning, and retrospective recall of task instructions, across the encoding and PM task type conditions.

4.1.1. *Determining an Appropriate Analysis Strategy*

The data for the variables in question included counts of infrequent events, and had little variance as these events were experimentally controlled. As such, they were not normally distributed. Non-parametric tests were therefore used throughout these preliminary analyses.

4.1.2. *The Occurrence of Errors across Encoding Conditions and PM Tasks*

The Errorful Learning conditions were designed to elicit at least two ‘forced’ errors, which were not present in the Errorless Learning conditions. Additionally, the encoding conditions varied in their inclusion of instructions to avoid versus encourage guessing, hence additional spontaneous errors could also occur. Two parallel error-rate analyses were therefore conducted, one including both forced and spontaneous errors (total error rate), and one excluding forced errors (spontaneous error rate). The actual error rates are given in Table 1.3.

For the Total Error rate, Friedman’s two-way Analysis of Variance by ranks for related samples confirmed that there were differences between the four different conditions ($p < .001$). Follow-up pairwise comparisons using Wilcoxon Signed Rank tests confirmed there were more errors in the Errorful compared with Errorless conditions, for both EBPM ($p < .001$) and TBPM ($p < .001$) tasks. The total error rates, however, did not differ between the EBPM and TBPM tasks overall ($p = .725$), nor between the two Errorful conditions ($p = .335$).

The equivalent analyses for Spontaneous Error rates showed a marginally significant difference across the four conditions (Friedman’s test, $p = .07$). Pairwise comparisons showed more errors in the Errorless than the Errorful condition (2 versus 14, $p = .027$). Such trends were also apparent within the EBPM conditions (0 versus 8, $p = .066$), and to a lesser extent the TBPM conditions (2 versus 5, $p = .180$). There was, importantly, no difference in spontaneous error rates between the EBPM and TBPM tasks (8 versus 7, $p = .725$).

These analyses show that the encoding manipulation was successful in increasing rates of forced (and, to a lesser extent, spontaneous) errors in the Errorful conditions compared with the Errorless conditions. As there were only two spontaneous errors in the Errorless conditions, compared with 14 in the Errorful conditions, this can also be taken as evidence that the paradigm was effective in reducing, if not completely eliminating, errors. Importantly, there was no indication that the rates of either total or spontaneous errors differed between the Event-based and Time-based PM tasks.

4.1.3. Exposure to Correct Information across Encoding Conditions and PM Tasks

When participants made erroneous responses during encoding, corrective feedback was immediately provided. This feedback took the form of either an easy question that elicited the correct response, or the direct provision of the correct information. Such information was also presented following “don’t know” responses. This increased the number of times that correct information was repeated, over and above the standard presentations within the instruction slides. To examine whether there were any differences in the exposure to correct information between the encoding conditions, a composite measure of ‘prompts plus spontaneous errors’ was created and used as the dependent variable in nonparametric analyses as above.

A Friedman test indicated that there was no difference in exposure rates between the four conditions ($p=.546$). Follow-up pairwise comparisons identified no differences between either of the EBPM conditions ($p=.396$) or the TBPM conditions ($p=.161$), nor between Errorless ($p=.655$) and Errorful conditions ($p=.826$).

These analyses indicate that there were no differences in overall exposure to the correct information between any of the conditions.

Table 1.3. Rates of errors and prompts by encoding condition and task type (M [SD]).

Condition / Measure	Errorless Encoding		Errorful Encoding	
	EBPM	TBPM	EBPM	TBPM
<i>Total Errors</i>	0 (0)	.143 (.363)	2.571 (1.016)	2.357 (.497)
<i>Spontaneous Errors</i>	0 (0)	.143 (.363)	.571 (1.016)	.357 (.497)
<i>Prompts</i>	.429 (.646)	.214 (.426)	.071 (.267)	.500 (.650)
<i>Exposure to correct information</i>	.429 (.646)	.357 (.497)	.643 (1.15)	.786 (.975)
<i>Retrospective Recall</i>	13 correct spontaneously, 1 after prompt	All correct spontaneously	All correct spontaneously	12 correct spontaneously, 2 after prompt

4.1.4. Retrospective Memory for PM Task Content

All participants were able to provide accurate information regarding PM and OT task content at the end of each testing condition. The majority reported this information in response to a general request to describe the task instructions. Two participants, however, required additional prompting to recall the precise details; one for the EF-TBPM condition, and one for both the EF-TBPM condition and the EL-EBPM condition. This prompting took the form of asking “you had another task to do too, to press a different button, which one was it? And when were you to press it?” These data were categorised into a binary ‘remembered unprompted’ versus ‘remembered with prompt’ variable, and entered into a Cochran’s Q test for related samples. The test was not significant ($p=.194$), suggesting that there were no systematic differences in retrospective memory recall between the conditions. Furthermore, this indicates that all participants retained the details of both the ongoing and PM tasks, and by implication, any PM task failures did not result from low-level failures of retrospective memory for task content.

4.2. Primary Analysis: The Impact of Encoding Condition on Prospective Memory Accuracy

This study’s primary hypothesis was that there would be an interaction between Encoding Condition and PM Task Type. Unfortunately with a final N of 14, the study was not sufficiently powered to detect interaction effects within a Repeated Measures ANOVA any smaller than $d = .8$. The analysis therefore prioritised the two contrasts of primary interest, firstly that of Event-based PM performance between the Errorless and Errorful Encoding conditions, and secondly of the Errorless Learning Advantage (ELA) for the EBPM task compared with the TBPM task, which is equivalent to testing the aforementioned interaction. These contrasts were assessed with paired t-tests, and effect size calculations. Parametric tests were used as they are sufficiently robust to deviations from the normal distribution to the extent that would be expected within a sample of this size, and with a range of scores from only 0-8 converted to proportions, no scores can really be considered to be outliers.

4.2.1. PM Accuracy as a Function of Encoding and Task Type

A one-tailed t-test for paired samples confirmed the hypothesised difference in Event-based PM performance between the two encoding conditions ($t(13)=2.274$, $p=.021$). After Errorful Encoding, 42.0% of PM targets were responded to correctly, whereas after Errorless Learning, 66.1% of targets were responded to correctly. The effect size Cohen’s d , calculated using Morris and DeShon’s (2002) method for dependent data, was .63, which is considered to be a medium effect (Cohen, 1992).

A paired t-test of the ELA difference scores, which is equivalent to testing the interaction between Encoding and task type did not reach the threshold for statistical significance, but it did indicate a trend in the predicted direction ($t(13)=1.514$, $p=.077$). Event-based PM performance was 24.1% better under Errorless Learning conditions compared with Errorful Learning conditions, whereas the equivalent Errorless Learning “Advantage” for Time-based PM performance was -.01%. Cohen’s d for this comparison was 0.41, considered small-medium. This pattern of results is illustrated in Figure 1.4.

With the benefit of hindsight it is clear that the final N of 14, though sufficient to detect the main effect of encoding on Event-based PM, offered insufficient statistical power to detect the small-medium sized interaction between encoding and PM task type.

Table 1.4. Descriptive statistics for PM performance and the Errorless Learning Advantage.

PM Type	Encoding Condition	Mean	Std. Dev	Std. Error
EBPM	Errorless	.661	.378	.101
	Errorful	.420	.472	.126
TBPM	Errorless	.615	.442	.118
	Errorful	.621	.476	.127
EBPM	ELA	.241	.397	.106
TBPM	ELA	-.006	.453	.121

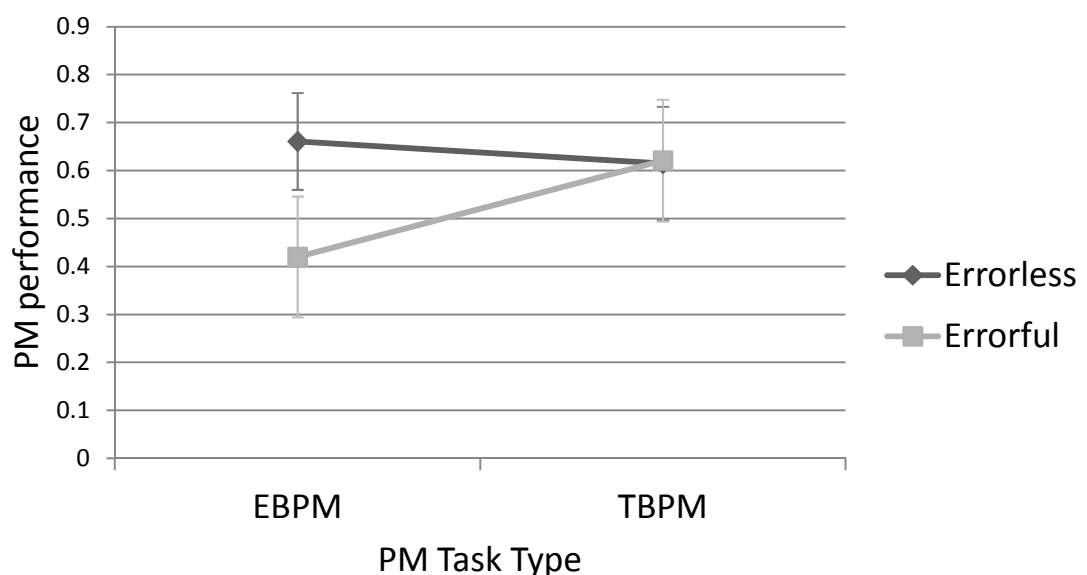


Figure 1.4. Mean EBPM and TBPM accuracy in Errorless and Errorful conditions.

Error bars show standard errors.

4.2.2. Time-based PM Errors by Encoding Condition

The nature of the Time-based task meant that in addition to ‘correct’ PM responses, incorrect PM responses could also be made¹⁶. A t-test for paired samples confirmed that there were no differences in the rate of PM commission errors between the two Encoding conditions ($t(13) = .165$, $p = .871$), with the mean number of commission errors being .71 (SD 1.90) in the Errorless condition, and .86 (SD 2.41) in the Errorful condition.

4.3. Ongoing Task Performance across the Experimental Conditions

In much the same way that increases in performance accuracy frequently occur at the detriment to speed of performance, it is conceivable that differences in performance of the PM task may occur at the expense of performance on the OT, reflecting for example a change in performance strategy. Given the difference in EBPM performance between the two Encoding conditions previously identified, it is important to examine any associated impact of Encoding or PM task type on OT performance.

4.3.1. Determining an Appropriate Analysis Strategy

The variables of Ongoing Task (OT) accuracy, number of trials, median RT and RTSD for each condition were inspected for normality using QQ plots, boxplots, and estimates of variance (standard deviations). All distributions approximated the normal distribution, and variance was approximately equal. As such, the data were deemed suitable for analysis within the General Linear Model. As there were no particular hypotheses regarding the effects of the experimental factors on OT measures, all four measures were included in a repeated measures Multivariate Analysis of Variance. This analysis allowed for the detection of ‘overarching’ effects that apply across some or all of the four dependent variables (i.e. the multivariate effects), in addition to effects on each individual dependent variable (i.e. the univariate effects).

4.3.2. Repeated Measures MANOVA on OT performance

A repeated measures Multivariate Analysis of Variance (MANOVA) was conducted on the ongoing task dependent variables of accuracy, number completed, median RT, and RTSD, with the within-subjects factors of Encoding (Errorless, Errorful), and PM Task (TBPM, EBPM).

The multivariate test showed no overall effect of Encoding ($F(4,10) = .171$, $p = .948$, $\eta^2 = .064$), nor of PM Task ($F(4,10) = 1.714$, $p = .223$, $\eta^2 = .407$), nor any interaction ($F(4,10) = 1.314$, $p = .329$, $\eta^2 = .345$). None of the associated Univariate tests were significant. The descriptive statistics are shown in Table 1.5.

¹⁶ Technically, it was also possible for PM errors of commission to occur within the Event-based task, but no such errors occurred.

Overall, this analysis suggests that the previously outlined effects of Encoding on PM accuracy were not at the detriment of OT performance.

Table 1.5. Descriptive statistics (M, [SEM], 95% Confidence Interval) for Ongoing Task variables according to experimental condition

	Errorless		Errorful	
	EBPM	TBPM	EBPM	TBPM
Accuracy	94.87% (1.61) <i>CI : 91.39 - 98.36</i>	96.13% (.65) <i>CI : 94.73 - 97.54</i>	95.84 (1.14) <i>CI : 93.38 - 98.30</i>	95.82% (.81) <i>CI : 94.06 - 97.57</i>
N Correct	258.36 (27.09) <i>CI : 199.84 - 316.87</i>	253.43 (25.64) <i>CI : 198.03 - 308.83</i>	260.72 (26.38) <i>CI : 203.74 - 317.69</i>	248.14 (25.52) <i>CI : 193.02 - 303.27</i>
Median	2411.75 (306.61) <i>CI : 1749.35 - 3074.15</i>	2611.68 (346.44) <i>CI : 1863.24-3360.12</i>	2502.46 (352.28) <i>CI : 1741.53 - 3263.40</i>	2694.36 (394.01) <i>CI : 1843.15 - 3545.56</i>
Correct RT				
Correct	1710.17 (308.89) <i>CI : 964.74 - 2230.25</i>	1490.76 (247.27) <i>CI : 956.56 - 2024.961</i>	1597.50 (292.89) <i>CI : 964.74 - 2230.25</i>	1556.05 (285.26) <i>CI : 939.80 - 2172.31</i>
RTSD				

4.4. PM Awareness Variables across Experimental Conditions

As stated, the ELA is thought to stem from increased reliance on implicit memory and/or through lesser reliance on explicit memory in screening out errors. If such mechanisms underlie the ELA, one might expect *reduced explicit awareness* of material learned errorlessly relative to that learned with errors. With this in mind, the next analysis examined participants' perceptions of their performance under the four conditions, using data from the post-task questionnaire.

4.4.1. Determining an Appropriate Analysis Strategy

The four awareness variables (PM Target Awareness, PM Response Awareness, Subjective Accuracy Estimate, and Subjective-Objective Accuracy Discrepancy) were inspected for normality using QQ plots, boxplots, and estimates of variance (standard deviations). There were three outliers in the PM response awareness variable, whose values were replaced by the mean plus two standard deviations (as recommended in Field, 2009). The distributions were otherwise approximately normal, and variance was approximately equal. As such, the data were deemed suitable for analysis within the GLM.

4.4.2. Repeated Measures MANOVA on Awareness variables

A repeated measures MANOVA was conducted on the dependent variables Target Awareness, Response Awareness, Subjective Accuracy Estimate, and Objective-subjective Accuracy

Discrepancy, with the factors Encoding (Errorless, Errorful) and PM Task Type (EBPM, TBPM). The Multivariate test showed a significant effect of Encoding ($F(4,10) = 4.363$, $p = .027$, $\eta^2 = .636$), however there was no main effect of PM Task Type ($F(4,10) = .670$, $p = .628$, $\eta^2 = .211$), and no interaction between Encoding and Task Type ($F(4,10) = .926$, $p = .486$, $\eta^2 = .270$).

The subsequent Univariate tests revealed a significant main effect of Encoding on PM Target Awareness ($F(1,13) = 10.591$, $p = .006$, $\eta^2 = .449$), such that post-test awareness of PM targets was reduced in the Errorless Encoding conditions (mean .616, 95% CI .438-.794) relative to the Errorful Encoding conditions (mean .897, 95% CI .723-1.072). There were, however, no significant effects of Encoding on PM Response Awareness ($F(1,13) = 1.910$, $p = .190$, $\eta^2 = .128$), Subjective Accuracy Estimate ($F(1,13) = .249$, $p = .626$, $\eta^2 = .019$) or the Subjective-Objective Discrepancy Score ($F(1,13) = .652$, $p = .434$, $\eta^2 = .048$).

If such awareness reductions stemmed from post-task suppression of target-related information, which could be viewed as the intention completion effect, one might expect there to be an inverse relationship between target awareness and PM performance, as each task is completed its accessibility is reduced. There was a marginally significant correlation between overall target awareness and overall PM performance ($r(14) = .428$, $p = .064$), but it was not in the predicted direction. Indeed, as can be seen from Figure 1.5, the effect of Encoding condition on Target Awareness was driven primarily by the reduced awareness in the Errorless Time-based task, where PM performance was actually equivalent between the two Encoding conditions.

Table 1.6. Descriptive Statistics for Awareness variables by Encoding condition

Measure	Encoding	Mean	Std. Error	95% Confidence Interval	
				Lower	Upper
PM Target Awareness	Errorless	.616	.082	.438	.794
	Errorful	.897	.081	.723	1.072
PM Response Awareness	Errorless	.939	.150	.614	1.264
	Errorful	.631	.099	.416	.846
Subjective Accuracy	Errorless	.842	.048	.738	.947
	Errorful	.801	.082	.624	.978
Subjective-Objective Discrepancy	Errorless	.205	.107	-.026	.436
	Errorful	.281	.111	.041	.521

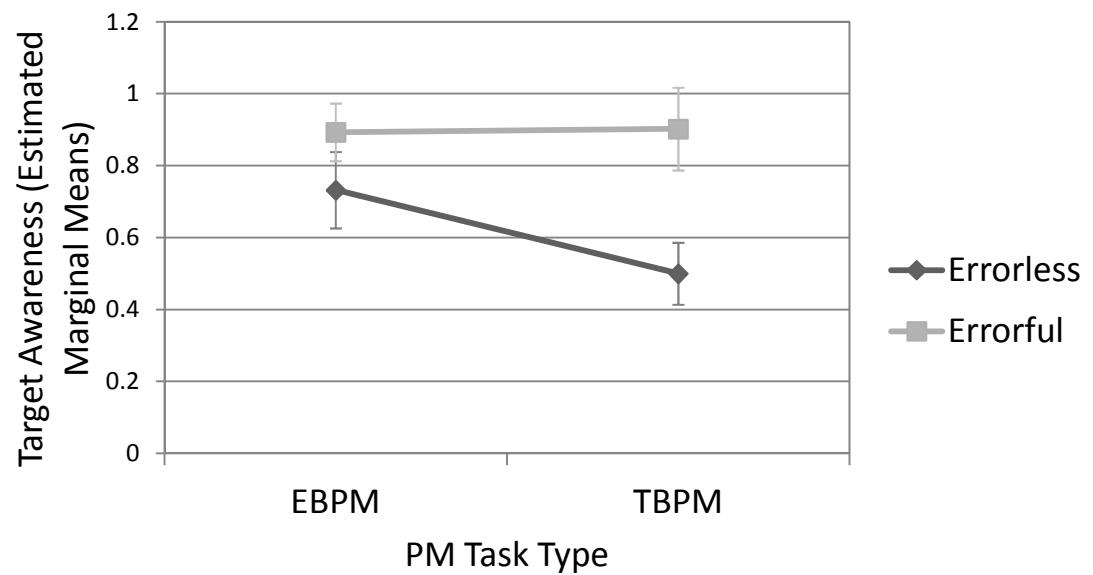


Figure 1.5. Mean post-test awareness of PM Targets plotted by Encoding Condition and PM Task Type.

Error bars show standard errors. Note that only the main effect of Encoding is significant.

5. Discussion

This final section sets out the main findings of the study, and links these findings to the published literature on PM and on EL, considering the implications of the experiment for our theoretical understanding of each. The limitations of the study are then discussed, along with a series of suggestions for further research. Finally, the clinical implications are described, with illustrative examples of how the findings could be applied to clinical work with people with memory disorders.

5.1. Summary of Findings

This study aimed to investigate the impact of errors made while encoding PM task instructions on subsequent performance of two different types of PM task; time based tasks, and Event-based tasks. One aspect of the encoding stage (the presence/absence of errors) was manipulated experimentally by asking participants to guess one part of the instruction before they could reasonably know what it was, and by encouraging them to guess if they were unsure of the correct response. An analysis of error rates during encoding showed that errors occurred at a significantly higher rate in the Errorful condition compared with the Errorless condition. This implies that the experimental manipulation was effective. Despite the difference in error rates, it was important that in both conditions the PM task was learned sufficiently well to allow subsequent retrieval, and that the Errorless condition was not associated with any increased exposure to the correct information than the Errorful condition. To this end, the correct information was presented 20 times per task, with additional presentations in the event of either errors or ‘don’t know’ responses. There were no differences in these rates across the conditions. Finally, post-test data showed that the intentions were adequately retained after a 20-minute delay, in all conditions. It therefore seems reasonable to conclude that the study was successful in its aim to manipulate error rates during encoding whilst maintaining equivalent exposure to correct information and ensuring adequate retrospective memory for the content of the PM task instructions.

Moving on to Prospective Memory performance, this study identified that, in a group of people with memory impairment, Event-based Prospective Memory performance was significantly better when task instructions had been encoded under Errorless Learning conditions than Errorful Learning conditions. This “Errorless Learning Advantage” (ELA) had a moderate effect size ($d=.63$). In contrast, there was no ELA for Time-based tasks. The interaction between errors during encoding and PM task type was not statistically significant in this small sample, but there was a clear trend, and the effect was of a small-moderate size ($d=.41$). Furthermore, this effect occurred without any detrimental effect on Ongoing Task

performance, which may have been expected if participants had, for example, intentionally prioritised the PM task in the Errorless Learning conditions.

The present finding of a beneficial effect of EL on EBPM is consistent with earlier studies, and it extends them by showing that this technique known to be effective for memory can have subsequent effects on behaviour. The finding that there was no ELA for TBPM is also in line with the hypotheses, but is rather more difficult to interpret, as null results can of course stem from a lack of statistical power, or use of an insensitive task. However, there are several observations that speak against this explanation. The EB and TB tasks were balanced in the number of actions required, and in for the number of words in the instruction sentence. Their demands in terms of encoding (i.e. information to be remembered) and performance (i.e. the response required, the number of responses required) were therefore very similar. Furthermore, when EB and TB PM scores were collapsed across encoding conditions, accuracy was clearly at an equivalent level, and with very similar variance. Therefore, it does not seem likely that any simple measurement confound precluded detection of an ELA within the TBPM task. This, along with the near-complete overlap in TBPM scores in the Errorless and Errorful Learning conditions, suggests at the very least that if an ELA could be detected for the TBPM task given a sufficiently large sample, then it would be significantly smaller than the ELA for EBPM.

An additional exploratory analysis of metacognitive awareness of PM task-related information identified that participants' awareness of PM targets was reduced in the Errorless compared with the Errorful Encoding condition. There was no such effect on participants' awareness of their own PM responses, or overall estimates of their accuracy, or in the accuracy of those estimates. This finding is considered below.

5.2. Relation of the Findings to the Literature

In this section, the findings are discussed in relation to the models of PM outlined in the introduction, the body of research on Errorless Learning, and in relation to the rationale for examining the impact of EL on EBPM and TBPM.

5.2.1. The Results in Relation to Models of Prospective Memory

PM tasks are composed of a series of stages, from encoding through retrieval to performance and evaluation. This study focussed on experimental manipulation of a factor at the encoding stage to examine any later impact at the retrieval stage. This approach is frequently taken in the PM literature, where data from participants who do not remember the PM task instructions at the end of the test are often excluded from subsequent analyses, as it is assumed that this failure to report the task instructions represents a task failure at the lowest

level of the hierarchy. This approach was also adopted in the current study. However, it is interesting to consider the potential effects of EL approaches at other levels of the hierarchy. Failures at these lower levels are likely very important in determining success or failure on PM tasks in everyday life, and if PM task instructions are more likely to be stored after Errorless than Errorful encoding procedures, then of course they would be more likely to be subsequently acted upon (i.e. if a task is not learned it cannot later be performed). The question that this study aimed to address was whether there would be an impact of errors made during learning at higher-level stages of the PM model, specifically at retrieval and the action initiation stages. The primary result, that PM performance was better after Errorless Learning than Errorful Learning for the Event-based task, supports the idea that differences in the encoding stage can have benefits at higher-level stages of retrieval and/or action initiation. It may well be the case that in more naturalistic settings, without the need to ensure adequate post-test retrieval across experimental conditions, the benefits of EL would be even larger.

The observed trend towards an interaction between encoding condition and task type also supports multi-process accounts of prospective memory. If the same retrieval processes were required in both TBPM and EPBM, then no such differential effect would be expected. However, this is not to say that the result provides any strong evidence against the competing Preparatory Attention and Memory model, as there were no “no PM task” conditions that would speak to the issue of whether simply holding an intention in mind is sufficient to reduce Ongoing task performance.

5.2.2. The Results in Relation to Research on Errorless Learning

As set out in the introduction, EL has consistently been found to have beneficial effects on learning verbal information and face-name associations. In their paper from 2000, Evans et al. stated that EL was likely to show benefits relative to Errorful Learning in conditions that facilitate implicit retrieval, giving the examples that Errorless methods would facilitate learning of face-name associations with a first-letter cue, whereas it would not facilitate explicit recall of novel associations, such as those involved in route learning or programming an electronic organiser. Several other studies have also concluded that the ELA is mediated by implicit rather than explicit memory processes (e.g. Page et al., 2006, Anderson & Craik, 2006). This seems like a valid argument, but there are some caveats. Errorless Learning *has* been found to improve performance on the two tasks that Evans et al. hypothesised it would not; programming an electronic organiser in an earlier case study by Wilson et al. (1994), and learning a novel route in a subsequent study by Lloyd et al. (2009). A range of factors related to the experimental designs may have contributed to these discrepancies, for example, differences in the patient groups studied (e.g. severity of memory deficit, presence of other

cognitive impairments), and in the implementation of the learning procedures (e.g. if the Errorless condition was particularly passive or involved reduced opportunity for retrieval practice, a detrimental effect on performance may be expected). It is also conceivable that performance on route learning tasks may be influenced by implicit memory, either through facilitation of the correct responses or interference from prior errors, so perhaps the implicit/explicit distinction holds even if the specific examples previously given do not. Relating these existing findings to those from the current study, it could be argued that the retrieval process involved in the EBPM task is more analogous to the 'implicit' conditions referred to by Evans et al. (as a previously-learned external cue is presented for the participant to act upon), and the retrieval process involved in the TBPM task is more akin to the 'explicit' conditions referred to (as the participant learns a task instruction but has to retrieve of their own accord). So far, the present results appear to tally with the literature.

Recent neuroimaging studies in healthy volunteers have identified that the ELA is associated with reduced PFC activation compared with Errorful Learning (Hammer et al., 2013), and that temporary disruption of left PFC functioning impairs Errorful Learning but not Errorless Learning (Hammer et al., 2011). This is consistent with the broader neuroimaging literature on memory, which shows that priming (implicit memory) is associated with decreased BOLD signal in frontal, temporal and occipital cortex (Henson, 2003), whereas recognition memory (a form of explicit memory) is associated with increased BOLD signal in prefrontal, parietal and medial temporal cortex (Henson, 2005). If the ELA stems from circumnavigating the frontal demands of errorful memory retrieval, it follows that provided other factors (e.g. presentation time, elaboration, retrieval practice) are held constant, there should be an ELA in *any* memory retrieval task, not merely those where implicit memory is thought to be important for task performance. The present results are not in keeping with this prediction, as the TBPM task involved memory retrieval, and showed no benefit from EL. However, we should keep in mind that the neuroimaging studies to date have only examined cued recall tasks, and future research will likely qualify this prediction.

Given these neuroimaging results, would an ELA also be expected for executive or attentional tasks? It is conceivable that as Errorless Learning reduces the frontal demands of memory retrieval tasks, there would be a certain "freeing up" of frontal resources to be allocated to more purely attentional or executive demands. However, the frontal structures involved in episodic memory retrieval would not necessarily overlap with those involved in attentional or executive tasks, and not with PM tasks, which are most closely associated with anterior PFC. Therefore, it seems sensible to conclude that EL would be unlikely to hold benefits for

performance of executive or attentional tasks, and indeed such effects were not observed in the present study.

5.2.3. The Results in Relation to Research on Metacognition

The study generated an unexpected finding that Errorless Learning was associated with reduced awareness of PM targets. The counterbalancing of the experimental conditions means that this is not a simple effect of order (e.g. target awareness estimates get more accurate with practice), and the within-subjects design means that such differences cannot be accounted for by individual differences in memory functioning, or in estimation ability. Furthermore, the effect was only present for the measure of Target Awareness (how many chances to press the PM key were there?), not the putatively very similar measure of Response Awareness (how many times did you press the PM key?). Furthermore, this was a main effect of Encoding, and therefore applied equally to the Time-based and Event-based PM Tasks. There is a small chance that this is a spurious finding, but as this aspect of the study was exploratory rather than underpinned by strong hypotheses, it would be remiss not to consider it carefully.

Within the PM tasks used in this study, online awareness must have involved several sub-processes. These may have included perceptual and semantic processing of the target within the Ongoing task, and the linking of that target to the stored representation within the PM task instruction. If the target was linked to the stored instruction in time to make the appropriate action, it would have been registered as a completed task. If the target was linked/recognised too late, after initiation of the Ongoing task action, it would have been experienced as an error. If the target was not linked to the stored intention at all, it is unlikely that it would have been processed as an error, though the process of answering the post-test questions could in principle provoke a delayed realisation that errors were made. These online processes likely contributed towards Target Awareness responses, through their building an overall sense of what happened during the task. However, Target Awareness responses were very much “offline” estimates, and as such involved a significant episodic memory component also, in reconstructing the task context to provide an answer to the question. The online awareness processes may also be subject to individual biases (e.g. errors may be more emotionally salient and hence remembered better than correct responses for some, and for others the reverse may be true), which may add a degree of noise to the offline estimate.

The Errorless/Errorful encoding manipulations were shown to have an effect on Target Awareness, and given the previous consideration of the processes that contribute to target awareness, it is conceivable that the simplified task representation built up through Errorless Learning resulted in a type of specific automatized association between target times, events

and actions, making targets that were not recognised and acted upon less salient, and reducing Target Awareness (and it is interesting to note that awareness of correct PM *responses* was near-perfect after EL, at 94%, in comparison with 63% after Errorful Learning, though this difference was not statistically significant). The more contextualised Errorful Learning procedure and its resultant broader PM task representation could conversely have been seen to lead to a less specific matching process that served to increase the awareness of targets to which they had made an incorrect response. This effect clearly requires replication, but it raises interesting questions nonetheless.

5.3. Limitations of the Study, and Suggested Modifications in Future Research

5.3.1. Sample Size and Associated Shortcomings

This study has a number of limitations. Firstly, it had limited statistical power to detect small or medium-sized effects. On the basis of the effect size of the trend-level interaction observed in this study, 38 participants would be required in any future study to have 80% power to detect the difference. This would clearly not be possible within the time frame that was available to complete the project. If further recruitment sites had been approached then this may have increased numbers somewhat, but unlikely to the required extent within the required time-frame. A related limitation is that due to lower-than-anticipated recruitment rates, the Event-based tasks were completed prior to the Time-based tasks in 8 cases (and only 6 the reverse). However, it is difficult to imagine that this would have had any clear impact on the results obtained. Finally, the initial protocol specified that Errorless and Errorful Learning conditions should be separated by one week, to minimise the likelihood that instructions to avoid/encourage errors may contaminate other conditions (i.e. to avoid carry-over effects). As it was quite challenging to recruit even this number of participants, in six cases where participants preferred not to have two appointments, all conditions were completed in one session. The analysis of spontaneous error rates during encoding suggests, however, that this deviation from protocol did not result in significant contamination of the two learning conditions, and this is something useful to know when planning subsequent research projects.

Despite the above limitations, it is also important to note that even though this study is preliminary in nature, the results and in particular the effect sizes generated are informative and provide a good basis for subsequent work. The sample size is also comparable with many other similar studies. For example, there were 20 people with memory impairment in Lloyd et al.'s (2009) study, 19 people with amnesic MCI in Lubinsky et al.'s (2009) study, 16 people with severe memory impairment in Baddeley and Wilson's (1994) study, 14 and 16 people with Alzheimer's disease respectively in the two experiments reported by Kinsella et al. (2007), 12

people with memory impairment in Grilli & McFarland's (2011) study, and just ten people with AD in Dunn and Clare's (2007) study, all of which were experiments that examined the effects of different encoding methods on subsequent memory performance.

5.3.2. Limitations Related to Measurement Decisions

Pre-existing test scores were used to define eligibility for the study, and these tests were not the same for all participants. This was a pragmatic choice made in the interests of aiding the identification of eligible participants, and minimising the demands of the participation for potential volunteers. It was also effective in the setting in which the research was conducted, as recent test scores were available for the vast majority of potential participants. However, in an ideal world, participants would have completed a novel battery of tests to measure more precisely the current memory impairment, and performance in other cognitive domains.

The measure of awareness used in this study likely had rather a strong loading on episodic (retrospective) memory, as well as general intellectual ability as it required making cognitive estimates on the basis of the task they had just completed. This is just one "off-line" aspect of metacognitive awareness. As noted in the introduction, some studies have also asked participants to predict PM performance, and this is one approach that could have been taken here. Certainly it would have been interesting to see if there were any differences in the accuracy of predictions compared with retrospective estimates (c.f. S. Smith et al.'s finding that people with Parkinson's Disease are impaired at making predictions but not postdictions). It may also be interesting to measure "on-line" awareness, for example by ask people to report when they noticed a PM error or correct PM response, an approach taken by McAvinue, O'Keeffe, McMackin and Robertson (2005) in a sustained attention paradigm. However, such approaches may have unwanted effects on performance; for example increasing error awareness often facilitates task performance. As such, future studies of metacognitive awareness of PM performance may wish to use indirect autonomic measures such as skin conductance response, which has again been recently demonstrated in the field of sustained attention (e.g. Hoerold, Pender, & Robertson, 2013).

Another measurement issue that could be considered a limitation of the current study is the use of a rather artificial laboratory task to measure PM performance. This was a necessary first step in examining the impact of errors during learning on PM performance, and ideally future studies would improve upon this aspect of the design by measuring performance on more naturalistic tasks.

Finally, the use of Time-based and Event-based PM tasks in this study is a potential source of criticism. These tasks were selected as they are thought to involve different underlying

retrieval processes, with Time-based tasks having more frontal attentional/executive demands, and Event-based tasks having more purely mnemonic retrieval processes. As discussed previously, the distinctions between these tasks are not necessarily so clear cut, for example Time-based tasks also have clear mnemonic demands. One suggestion for how to address this issue in future would be to run the experiment again, but rather than comparing TB and EB tasks, to compare EBPM tasks with either focal or nonfocal retrieval cues. Here the task instructions could be very similar (focal: press red when you see the word car; nonfocal: press red when you see the letters “mor”), but the retrieval demands are rather different, with the nonfocal cues requiring more attentional monitoring than focal cues (Scullin et al., 2010). It may even be possible to use exactly the same instructions (e.g. press red when you see the letters “for”), but simply to alter the position of the retrieval cue such that it appears in either the focal initial position, as in “fortunate”, or in a non-focal later position, as in “misfortune”. Though this sort of design would not address the issue that although both tasks have mnemonic demands, only one has substantial attentional demands, it would mean that the task instructions were very closely matched, whilst still manipulating the degree to which attentional monitoring is required for successful task performance. It is quite difficult to imagine a set of PM tasks that have orthogonal attentional and mnemonic demands, but it would in principle be possible to equate the overall attentional demands of the tasks across the PM paradigm by making the OT more attentionally demanding in the focal PM condition than the nonfocal PM condition¹⁷. Any remaining effects of PM cue type would then suggest very strongly that there were different retrieval processes underlying focal and nonfocal PM performance.

5.4. Future Directions

There are many potential future studies to be conducted in this area. In the initial literature review it was suggested that patient fMRI studies of Errorless Learning, along with general fMRI studies comparing TBPM and EBPM may help to address some inconsistencies in the literature. An updated meta-analysis of the effects of EL on memory would also be useful. More research is also needed on the validity and reliability of various means of assessing PM. In the preceding limitations section, modifications and further studies were also recommended, including obtaining larger samples, incorporating on-line and indirect measures of awareness, and more precisely matching the demands of the PM tasks. This section will focus, however, on future studies that follow more specifically from the experimental results.

¹⁷ For example by means of manipulating parameters of Bundesen’s Theory of Visual Attention tasks (Bundesen, 1990; Finke et al., 2005).

5.4.1. Studying Errorless Learning of Naturalistic Prospective Memory Tasks

The assessment approach used in this study is acknowledged to be lacking in ecological validity. Three factors influenced the choice of outcome measures: the limited time available (one day per week), the scarcity of potential participants, and the intricacy of the experimental questions. In light of these, it seemed more sensible to use laboratory tasks, as they allowed for the accurate measurement of PM and ongoing task activity, with plenty of trials, in a relatively brief time duration. However, PM is very amenable to measurement in an ecologically valid manner, and if the current findings are to be translated into effective strategies for use in PM rehabilitation, it would be important to verify the impact of EL on naturalistic PM task performance. One potential study would be to use an ABAB experimentally designed case series whereby participants learn the instructions to TBPM tasks (e.g. to make phone calls at set times), and in response to particular events (e.g. receiving a text message from the research team), over week-long phases with Errorless Learning one week, and errorful the next. Conversely, single-case studies could involve the identification and categorisation of everyday PM tasks, and use Errorless Learning to teach one subset of tasks, and Errorful methods to teach the other. Another alternative would be to replicate the experiment using a virtual reality prospective memory task.

5.4.2. Studying Naturalistic Encoding Errors

The error manipulation used in this study was also rather artificial. The majority of the errors made during the encoding phase were provoked by the experimenter, on one occasion for each task case this involved the experimenter providing the incorrect information, and on the other occasion, by asking participants to guess the intended results. Additional unprompted errors were made, but not by all subjects. This is clearly not a naturalistic study of the impact of errors on learning, and it is interesting to consider the differential impact of provoked versus spontaneous errors on performance.

Most studies of Errorless Learning artificially provoke errors to ensure an effective comparison with Errorless Learning conditions (though there are exceptions, e.g. Dunn & Clare, 2007; Lloyd et al., 2009). Some studies have however manipulated the types of errors produced, and these studies are informative in the consideration of this issue of how different types of error may affect subsequent learning and behaviour. Lubinsky, Rich and Anderson (2009) examined whether self-generation of material during learning altered the ELA in a verbal paired associate learning task. They found evidence that it did enhance the ELA, at least where there was a high degree of study-test similarity. Moreover, this study afforded the opportunity to examine priming effects of errors that had been self- versus experimenter-generated, and indeed there was increased priming for self-generated errors. This suggests that there is increased

interference from self-generated errors than experimenter-generated errors. Though both types of error were provoked by the experimenter, rather than being truly naturalistic/unprovoked errors, if one were to extrapolate from these findings, it would lead to the prediction that naturally occurring errors would interfere more than provoked errors, and hence the advantage for errorlessly learned material would be even greater when compared with more naturalistic errors. One could imagine several factors that may make naturalistic errors more ‘captivating’ than provoked errors, for example the context surrounding their production would be more similar to that surrounding the production of correct responses than that surrounding provoked errors (e.g. both correct responses and naturalistic errors may be provoked by the same question, such as “can you remember the word that started with B”, whereas provoked errors in response to a rather odd question such as “guess the word on this card, it starts with B”, before they have even seen it). Further, it is likely that naturalistic errors are produced with the aim of producing the correct response, rather than simply a guess. Also, it may be the case that the errors produced naturalistically are more salient than provoked errors, for example in their association with existing knowledge. Future studies could examine the impact of these different error types on subsequent memory performance.

5.4.3. Disseminating the Findings and Practical Applications of the Results

These results indicate that Errorless Learning methods may be helpful in promoting action in addition to improving learning. This is quite an important finding, as Prospective Memory tasks are part and parcel of everyday life, and are often very challenging for people with cognitive impairment. Many rehabilitation goals directly concern PM tasks (e.g. Peter will remember to take his phone with him every time he leaves the house, Anne will remember to take her medication in the mornings), and many further will have PM components (e.g. Sarah will learn to manage her finances independently, Dylan will do his grocery shopping on a weekly basis). There are various possibilities regarding how EL methods could be used clinically to improve PM performance. For example:

- a) Facilitating performance on specific PM tasks: For Event-based PM tasks that are not reliably completed, distil them to their simplest instruction, and teach these instructions using Errorless methods. A selection of examples may be: When I clean my teeth I will remember to take my tablets; Before I go upstairs to bed I will lock the front door; When I have something to eat, I will also drink a glass of water; As I leave the house, I will say to myself “have I got my keys?”.
- b) Establishing new behaviours, and replacing unwanted behaviours: specify and teach a clear goal e.g. When I finish dinner I will have an apple for dessert.

- c) Enhancing performance on procedural tasks with PM components: when focussing on Activities of Daily Living in rehabilitation, prompting and fading of instructions over time are well-known effective approaches, but the inclusion of error minimisation techniques within these approaches may improve learning further and/or facilitate acting upon PM sub-goals.
- d) Promoting effective interactions with people with memory impairment: It is always helpful to remember that using error-minimising processes (i.e. communicating clearly), perhaps particularly when giving instructions but also in more general interactions, will likely have better outcomes than approaches that incorporate errors (e.g. giving confused or revised instructions). This sort of reminder may be particularly helpful to staff and family members unused to being with people with memory impairment.

Finally, I think it would be useful to disseminate the findings of this study and the broader review of the EL literature. In the literature, EL has tended to be characterised as a passive approach and the original Baddeley and Wilson (1994) study specified as a rehabilitation technique, whereas it was designed as a proof-of-concept study (which was followed the same year with a paper reporting several case studies illustrating several the clinical applications of the technique). Consequentially, some implementations of Errorless Learning focus on simply giving the answers to question rather than helping to elicit the correct information and avoid errors. To ensure that the technique is used effectively, dissemination of findings through research papers and conference presentations are needed to enhance correct understanding of how Errorless Learning techniques have been developed; within these, it will be advisable to include a very brief statement describing the study's clinical implications or applications in the abstract of any publications, rather than stating simply that such implications exist, or that they are discussed in the main paper. It may also be worthwhile to write a simple leaflet aimed at service users, relatives and support staff which includes "take-home messages" and concrete examples of the clinical use of the technique.

6. References

- Adda, C. C., Castro, L. H. M., Além-Mar e Silva, L. C., de Manreza, M. L. G., & Kashiara, R. (2008). Prospective memory and mesial temporal epilepsy associated with hippocampal sclerosis. *Neuropsychologia*, 46(7), 1954–1964. doi:10.1016/j.neuropsychologia.2008.01.016
- Agnew, S. K., & Morris, R. G. (1998). The heterogeneity of anosognosia for memory impairment in Alzheimer's disease: A review of the literature and a proposed model. *Aging & Mental Health*, 2(1), 7–19.
- Alexander, M. P., Martin, M., McDaniel, M. A., & Einstein, G. O. (2004). Importance effects on performance in event-based prospective memory tasks. *Memory*, 12(5), 553–561.
- Anderson, N. D., & Craik, F. I. M. (2006). The mnemonic mechanisms of errorless learning. *Neuropsychologia*, 44(14), 2806–2813. doi:10.1016/j.neuropsychologia.2006.05.026
- Anderson, N. D., Guild, E. B., Cyr, A.-A., Roberts, J., & Clare, L. (2012). Contributions of frontal and medial temporal lobe functioning to the errorless learning advantage. *Neuropsychological Rehabilitation*, 22(2), 169–186. doi:10.1080/09602011.2011.639609
- Baddeley, A. D. (1997). *Human Memory: Theory and Practice*. Hove: Psychology Press.
- Baddeley, A. D., Emslie, H., & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language Processing Test*. Thames Valley Test Company Bury St Edmunds.
- Baddeley, A. D., Nimmo-Smith, I., & Emslie, H. (1994). *Doors and People*. Bury St Edmunds: Thames Valley Test Company.
- Baddeley, A. D., & Wilson, B. A. (1994). When implicit learning fails: amnesia and the problem of error elimination. *Neuropsychologia*, 32(1), 53–68.
- Bier, N., Van Der Linden, M., Gagnon, L., Desrosiers, J., Adam, S., Louveaux, S., & Saint-Mleux, J. (2008). Face–name association learning in early Alzheimer's disease: A comparison of learning methods and their underlying mechanisms. *Neuropsychological Rehabilitation*, 18(3), 343–371. doi:10.1080/09602010701694723
- Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical Aspects of Memory: Current Research and Issues, Vol. 2: Clinical and Educational Implications* (pp. 283–288). Chichester: John Wiley & Sons.
- Brandimonte, M. A., Filippello, P., Coluccia, E., Altgassen, M., & Kliegel, M. (2011). To do or not to do? Prospective memory versus response inhibition in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Memory*, 19(1), 56–66. doi:10.1080/09658211.2010.535657

- Brandimonte, M. A., & Passolunghi, M. C. (1994). The effect of cue-familiarity, cue-distinctiveness, and retention interval on prospective remembering. *The Quarterly Journal of Experimental Psychology*, 47(3), 565–587.
- Brooks, N., Campsie, L., Symington, C., Beattie, A., & McKinlay, W. (1987). The effects of severe head injury on patient and relative within several years of injury. *Journal of Head Trauma Rehabilitation*, 2, 1–13.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97(4), 523–547.
- Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M.-A., Dawson, D. R., ... Channon, S. (2006). The case for the development and use of “ecologically valid” measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, 12(2), 194–209. doi:10.1017/S1355617706060310
- Burgess, P. W., Gonen-Yaacovi, G., & Volle, E. (2011). Functional neuroimaging studies of prospective memory: what have we learnt so far? *Neuropsychologia*, 49(8), 2246–2257. doi:10.1016/j.neuropsychologia.2011.02.014
- Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, 38(6), 848–863.
- Carey, C. L., Woods, S. P., Rippeth, J. D., Heaton, R. K., & Grant, I. (2006). Prospective memory in HIV-1 infection. *Journal of Clinical and Experimental Neuropsychology*, 28(4), 536–548. doi:10.1080/13803390590949494
- Chan, R. C. K., Lui, S. S. Y., Wang, Y., Liu, A. C. Y., Chui, W. W. H., Shum, D. H. K., & Cheung, E. F. C. (2012). Patients with bipolar disorders share similar but attenuated prospective memory impairments with patients with schizophrenia. *Psychological Medicine*, 1–11. doi:10.1017/S003329171200236X
- Chan, R. C. K., Wang, Y., Ma, Z., Hong, X., Yuan, Y., Yu, X., ... Gong, Q. (2008). Objective measures of prospective memory do not correlate with subjective complaints in schizophrenia. *Schizophrenia Research*, 103(1–3), 229–239. doi:10.1016/j.schres.2008.02.019
- Clare, L., & Jones, R. S. P. (2008). Errorless learning in the rehabilitation of memory impairment: a critical review. *Neuropsychology Review*, 18(1), 1–23. doi:10.1007/s11065-008-9051-4
- Clare, L., Wilson, B. A., Carter, G., Breen, K., Gosses, A., & Hodges, J. R. (2000). Intervening with everyday memory problems in dementia of Alzheimer type: an errorless learning approach. *Journal of Clinical and Experimental Neuropsychology*, 22(1), 132–46. doi:10649552

- Clare, L., Wilson, B. A., Carter, G., Hodges, J. R., & Adams, M. (2001). Long-term maintenance of treatment gains following a cognitive rehabilitation intervention in early dementia of Alzheimer type: A single case study. *Neuropsychological Rehabilitation*, 11(3), 477–494.
- Cockburn, J. (1995). Task interruption in prospective memory: A frontal lobe function? *Cortex*, 31(1), 87–97.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Coughlan, A. K., Oddy, M., & Crawford, J. R. (2007). *BIRT Memory and Information Processing Battery (BMIPB)*. London: Brain Injury Rehabilitation Trust.
- Crawford, J. R., Henry, J. D., Crombie, C., & Taylor, E. P. (2001). Normative data for the HADS from a large non-clinical sample. *The British Journal of Clinical Psychology*, 40(Pt 4), 429–434.
- Crawford, J. R., Smith, G., Maylor, E. A., Della Sala, S., & Logie, R. (2003). The Prospective and Retrospective Memory Questionnaire PRMQ: Normative data and latent structure in a large non-clinical sample. *Memory*, 11(3), 261–275.
- Cuttler, C., & Graf, P. (2009). Sub-clinical compulsive checkers show impaired performance on habitual, event- and time-cued episodic prospective memory tasks. *Journal of Anxiety Disorders*, 23(6), 813–823. doi:10.1016/j.janxdis.2009.03.006
- Cyr, A.-A., & Anderson, N. D. (2012). Trial-and-error learning improves source memory among young and older adults. *Psychology and Aging*, 27(2), 429–439. doi:10.1037/a0025115
- D' Ydewalle, G., Bouckaert, D., & Brunfaut, E. (2001). Age-related differences and complexity of ongoing activities in time- and event-based prospective memory. *The American Journal of Psychology*, 114(3), 411–423.
- Dawkins, N., Cloherty, M. E., Gracey, F., & Evans, J. J. (2006). The factor structure of the Hospital Anxiety and Depression Scale in acquired brain injury. *Brain Injury*, 20(12), 1235–1239. doi:10.1080/02699050601076414
- Doering, B. K., Conrad, N., Rief, W., & Exner, C. (2011). Living with acquired brain injury: self-concept as mediating variable in the adjustment process. *Neuropsychological Rehabilitation*, 21(1), 42–63. doi:10.1080/09602011.2010.525947
- Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cognitive Psychology*, 30(3), 257–303. doi:10.1006/cogp.1996.0008
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23(10), 475–483.

- Dunn, J., & Clare, L. (2007). Learning face–name associations in early-stage dementia: Comparing the effects of errorless learning and effortful processing. *Neuropsychological Rehabilitation*, 17(6), 735–754.
- Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In *Prospective Memory: Theory and Applications*.
- Einstein, G. O., & McDaniel, M. A. (2010). Prospective memory and what costs do not reveal about retrieval processes: A commentary on Smith, Hunt, McVay, and McConnell (2007). *Journal of experimental psychology. Learning, memory, and cognition*, 36(4), 1082–1088; discussion 1089–1095. doi:10.1037/a0019184
- Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General*, 134(3), 327.
- Ellis, J. A. (1988). Memory for future intentions: Investigating pulses and steps. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical Aspects of Memory: Current Research and Issues* (Vols. 1-2, Vol. 1, pp. 371–376). Chi: Wiley.
- Ellis, J. A. (1996). Prospective memory or the realization of delayed intentions: A conceptual framework for research. *Prospective memory: Theory and applications*, 1–22.
- Ellis, J. A., & Freeman, J. E. (2008). Ten years on: Realizing delayed intentions. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective Memory: Cognitive, Neuroscience, Developmental, and Applied Perspectives* (pp. 1–27). New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates.
- Elvevag, B., Maylor, E. A., & Gilbert, A. L. (2003). Habitual prospective memory in schizophrenia. *BMC Psychiatry*, 3(9).
- Eslinger, P. J., & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology*, 35(12), 1731–1741.
- Evans, J. J., Wilson, B. A., Schuri, U., Andrade, J., Baddeley, A. D., Bruna, O., ... Laaksonen, R. (2000). A comparison of errorless and trial-and-error learning methods for teaching individuals with acquired memory deficits. *Neuropsychological Rehabilitation*, 10(1), 67–101.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Field, A. (2009). *Discovering Statistics Using SPSS* (Third Edition.). SAGE Publications Ltd.

- Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2006). The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation*, 16(2), 129–54. doi:Q71H14PV7M1815M1
- Fillingham, J. K., Sage, K., & Lambon-Ralph, M. A. (2005). Further explorations and an overview of errorless and errorful therapy for aphasic word-finding difficulties: The number of naming attempts during therapy affects outcome. *Aphasiology*, 19(7), 597–614.
- Finke, K., Bublak, P., Krummenacher, J., Kyllingsbaek, S., Muller, H. J., & Schneider, W. X. (2005). Usability of a theory of visual attention (TVA) for parameter-based measurement of attention I: evidence from normal subjects. *Journal of the International Neuropsychological Society*, 11(7), 832–842.
- Fish, J., Evans, J. J., Nimmo, M., Martin, E., Kersel, D., Bateman, A., ... Manly, T. (2007). Rehabilitation of executive dysfunction following brain injury: “content-free” cueing improves everyday prospective memory performance. *Neuropsychologia*, 45(6), 1318–30. doi:S0028-3932(06)00389-7
- Fish, J., Wilson, B. A., & Manly, T. (2010). The assessment and rehabilitation of prospective memory problems in people with neurological disorders: a review. *Neuropsychological Rehabilitation*, 20(2), 161–179. doi:10.1080/09602010903126029
- Fratalli, C. (2003). An errorless learning approach to treating dysnomia in frontotemporal dementia. *Journal of Medical Speech-Language Pathology*, 12, 21–24.
- Fuermaier, A. B. M., Tucha, L., Koerts, J., Aschenbrenner, S., Westermann, C., Weisbrod, M., ... Tucha, O. (2013). Complex prospective memory in adults with attention deficit hyperactivity disorder. *PLOS One*, 8(3), e58338. doi:10.1371/journal.pone.0058338
- Gordon, B. A., Shelton, J. T., Bugg, J. M., McDaniel, M. A., & Head, D. (2011). Structural correlates of prospective memory. *Neuropsychologia*, 49(14), 3795–3800. doi:10.1016/j.neuropsychologia.2011.09.035
- Goschke, T., & Kuhl, J. (1993). Representation of intentions: Persisting activation in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(5), 1211.
- Grilli, M. D., & McFarland, C. P. (2011). Imagine that: self-imagination improves prospective memory in memory-impaired individuals with neurological damage. *Neuropsychological Rehabilitation*, 21(6), 847–859. doi:10.1080/09602011.2011.627263
- Groot, Y. C. T., Wilson, B. A., Evans, J. J., & Watson, P. (2002). Prospective memory functioning in people with and without brain injury. *Journal of the International Neuropsychological Society*, 8(05), 645–654.

- Hammer, A., Mohammadi, B., Schmicker, M., Saliger, S., & Munte, T. F. (2011). Errorless and errorful learning modulated by transcranial direct current stimulation. *BMC Neuroscience*, 12, 72. doi:10.1186/1471-2202-12-72
- Hammer, A., Tempelmann, C., & Münte, T. F. (2013). Recognition of face-name associations after errorless and errorful learning: an fMRI study. *BMC Neuroscience*, 14, 30. doi:10.1186/1471-2202-14-30
- Harrison, T. L., & Einstein, G. O. (2010). Prospective memory: are preparatory attentional processes necessary for a single focal cue? *Memory & Cognition*, 38(7), 860–867. doi:10.3758/MC.38.7.860
- Hart, T., & Evans, J. J. (2006). Self-regulation and Goal Theories in Brain Injury Rehabilitation. *The Journal of Head Trauma Rehabilitation*, 21(2), 142.
- Haslam, C., Bazen-Peters, C., & Wright, I. (2012). Errorless learning improves memory performance in children with acquired brain injury: a controlled comparison of standard and self-generation techniques. *Neuropsychological Rehabilitation*, 22(5), 697–715. doi:10.1080/09602011.2012.686820
- Haslam, C., Hodder, K. I., & Yates, P. J. (2011). Errorless learning and spaced retrieval: how do these methods fare in healthy and clinical populations? *Journal of clinical and experimental neuropsychology*, 33(4), 432–447. doi:10.1080/13803395.2010.533155
- Haslam, C., Moss, Z., & Hodder, K. (2010). Are two methods better than one? Evaluating the effectiveness of combining errorless learning with vanishing cues. *Journal of Clinical and Experimental Neuropsychology*, 32(9), 973–985.
- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Chichester: Wiley.
- Henson, R. N. A. (2003). Neuroimaging studies of priming. *Progress in neurobiology*, 70(1), 53–81.
- Henson, R. N. A. (2005). A mini-review of fMRI studies of human medial temporal lobe activity associated with recognition memory. *The Quarterly Journal of Experimental Psychology B, Comparative and Physiological Psychology*, 58(3-4), 340–360. doi:10.1080/02724990444000113
- Hodder, K., & Haslam, C. (2006, July 31). *Errorless learning: A comparison with other memory rehabilitation techniques*. Presented at the Satellite Symposium on Neuropsychological Rehabilitation, Triesenberg, Liechtenstein.
- Hoerold, D., Pender, N. P., & Robertson, I. H. (2013). Metacognitive and online error awareness deficits after prefrontal cortex lesions. *Neuropsychologia*, 51(3), 385–391. doi:10.1016/j.neuropsychologia.2012.11.019

- Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, 12(4), F9–15. doi:10.1111/j.1467-7687.2009.00848.x
- Hunkin, N. M., Squires, E. J., Parkin, A. J., & Tidy, J. A. (1998). Are the benefits of errorless learning dependent on implicit memory? *Neuropsychologia*, 36(1), 25–36.
- Huppert, F. A., Johnson, T., & Nickson, J. (2000). High prevalence of prospective memory impairment in the elderly and in early-stage dementia: Findings from a population-based study. *Applied Cognitive Psychology*, 14(7), S63–S81.
- Jasiewicz, J., Kearns, W., Craighead, J., Fozard, J. L., Scott, S., & McCarthy, J., Jr. (2011). Smart rehabilitation for the 21st century: the Tampa Smart Home for veterans with traumatic brain injury. *Journal of Rehabilitation Research and Development*, 48(8), vii–xviii.
- Jokel, R., & Anderson, N. D. (2012). Quest for the best: effects of errorless and active encoding on word re-learning in semantic dementia. *Neuropsychological Rehabilitation*, 22(2), 187–214. doi:10.1080/09602011.2011.639626
- Jones, R. S. P., & Eayrs, C. B. (1992). The use of errorless learning condition procedures in teaching people with a learning disability. *Mental Handicap Research*, 5, 304–312.
- Katai, S., Maruyama, T., Hashimoto, T., & Ikeda, S. (2003). Event based and time based prospective memory in Parkinson's disease. *British Medical Journal*, 74(6), 704–709.
- Kern, R. S., Green, M. F., Mintz, J., & Liberman, R. P. (2003). Does “errorless learning” compensate for neurocognitive impairments in the work rehabilitation of persons with schizophrenia? *Psychological Medicine*, 33(3), 433–442.
- Kern, R. S., Green, M. F., Mitchell, S., Kopelowicz, A., Mintz, J., & Liberman, R. P. (2005). Extensions of errorless learning for social problem-solving deficits in schizophrenia. *The American Journal of Psychiatry*, 162(3), 513–519. doi:10.1176/appi.ajp.162.3.513
- Kern, R. S., Liberman, R. P., Kopelowicz, A., Mintz, J., & Green, M. F. (2002). Applications of errorless learning for improving work performance in persons with schizophrenia. *The American Journal of Psychiatry*, 159(11), 1921–1926.
- Kerns, K. A., & Price, K. J. (2001). An investigation of prospective memory in children with ADHD. *Child Neuropsychology*, 7(3), 162–171. doi:10.1076/chin.7.3.162.8744
- Kervick, R. B., & Kaemingk, K. L. (2005). Cognitive appraisal accuracy moderates the relationship between injury severity and psychosocial outcomes in traumatic brain injury. *Brain Injury*, 19(11), 881–889. doi:10.1080/02699050400025273
- Kessels, R. P. C., & de Haan, E. H. F. (2003). Implicit Learning in Memory Rehabilitation: A Meta-Analysis on Errorless Learning and Vanishing Cues Methods. *Journal of Clinical and Experimental Neuropsychology*, 25(6), 805–814.

- Kim, H. J., Burke, D. T., Dowds, M. M., Boone, K. A., & Park, G. J. (2000). Electronic memory aids for outpatient brain injury: follow-up findings. *Brain Injury*, 14(2), 187–96.
- Kinsella, G. J., Mullaly, E., Rand, E., Ong, B., Burton, C., Price, S., ... Storey, E. (2009). Early intervention for mild cognitive impairment: a randomised controlled trial. *Journal of Neurology, Neurosurgery, and Psychiatry*, 80(7), 730–736. doi:10.1136/jnnp.2008.148346
- Kinsella, G. J., Murtagh, D., Landry, A., Homfray, K., Hammond, M., O’Beirne, L., ... Ponsford, J. (1996). Everyday memory following traumatic brain injury. *Brain Injury*, 10(7), 499–508.
- Kinsella, G. J., Ong, B., Storey, E., Wallace, J., & Hester, R. (2007). Elaborated spaced-retrieval and prospective memory in mild Alzheimer’s disease. *Neuropsychological Rehabilitation*, 17(6), 688–706.
- Kliegel, M., Jäger, T., & Phillips, L. H. (2008). Adult age differences in event-based prospective memory: A meta-analysis on the role of focal versus nonfocal cues. *Psychology and Aging*, 23(1), 203.
- Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2001). Varying the importance of a prospective memory task: Differential effects across time-and event-based prospective memory. *Memory*, 9(1), 1–11.
- Kliegel, M., Phillips, L. H., Lemke, U., & Kopp, U. A. (2005). Planning and realisation of complex intentions in patients with Parkinson’s disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 76(11), 1501–1505.
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., ... Westerberg, H. (2005). Computerized training of working memory in children with ADHD--a randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44(2), 177–186. doi:10.1097/00004583-200502000-00010
- Kodera, T. L., & Rilling, M. (1976). Procedural antecedents of behavioral contrast: a re-examination of errorless learning. *Journal of the Experimental Analysis of Behavior*, 25(1), 27–42. doi:10.1901/jeab.1976.25-27
- Kondel, T. K. (2002). Prospective memory and executive function in schizophrenia. *Brain and Cognition*, 48(2-3), 405.
- Kopp, B. (2012). A simple hypothesis of executive function. *Frontiers in Human Neuroscience*, 6, 159. doi:10.3389/fnhum.2012.00159
- Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. Morris, & R. N. Sykes (Eds.), *Practical Aspects of Memory* (pp. 625–632). London: Academic Press.

- Levine, B., Robertson, I. H., & Manly, T. (in preparation). The Goal Management Training Handbook.
- Lloyd, J., Riley, G. A., & Powell, T. E. (2009). Errorless learning of novel routes through a virtual town in people with acquired brain injury. *Neuropsychological Rehabilitation, 19*, 98–109. doi:10.1080/09602010802117392
- Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey. *American Psychologist, 57*(9), 705–717.
- Lubinsky, T., Rich, J. B., & Anderson, N. D. (2009). Errorless learning and elaborative self-generation in healthy older adults and individuals with amnesic mild cognitive impairment: mnemonic benefits and mechanisms. *Journal of the International Neuropsychological Society, 15*(5), 704–716. doi:10.1017/S1355617709990270
- Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic Books.
- Maylor, E. A. (1990). Age and prospective memory. *The Quarterly journal of experimental psychology. A, Human experimental psychology, 42*(3), 471.
- McAvinue, L., O'Keeffe, F., McMackin, D., & Robertson, I. H. (2005). Impaired sustained attention and error awareness in traumatic brain injury: implications for insight. *Neuropsychological Rehabilitation, 15*(5), 569–587. doi:10.1080/09602010443000119
- McClelland, J. L., Thomas, A. G., McCandliss, B. D., & Fiez, J. A. (1999). Understanding failures of learning: Hebbian learning, competition for representational space, and some preliminary experimental data. *Progress in Brain Research, 121*, 75–80.
- McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: a multiprocess framework. *Applied Cognitive Psychology, 14*(7), S127–S144. doi:10.1002/acp.775
- McDaniel, M. A., Glisky, E. L., Rubin, S. R., Guynn, M. J., & Routhieaux, B. C. (1999). Prospective memory: a neuropsychological study. *Neuropsychology, 13*(1), 103–110.
- McDaniel, M. A., & Scullin, M. K. (2010). Implementation intention encoding does not automatize prospective memory responding. *Memory & Cognition, 38*(2), 221–232. doi:10.3758/MC.38.2.221
- McKenna, P., & Warrington, E. K. (1983). *The Graded Naming Test*. Windsor, Berkshire: NFER-Nelson.
- McKittrick, L. A., Camp, C. J., & Black, W. (1992). Prospective memory intervention in Alzheimers Disease. *The Journal of Gerontology: Psychological Sciences, 47*, 337–343.
- Meacham, J. A., & Singer, J. (1977). Incentive effects in prospective remembering. *Journal of Psychology, 97*, 191–197.

- Meeks, J. T., Hicks, J. L., & Marsh, R. L. (2007). Metacognitive awareness of event-based prospective memory. *Consciousness and cognition*, 16(4), 997–1004. doi:10.1016/j.concog.2006.09.005
- Meier, B., von Wartburg, P., Matter, S., Rothen, N., & Reber, R. (2011). Performance predictions improve prospective memory and influence retrieval experience. *Canadian Journal of Experimental Psychology*, 65(1), 12–18. doi:10.1037/a0022784
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49(2), 270–291. doi:10.1037/a0028228
- Michie, S., & Johnston, M. (2004). Changing clinical behaviour by making guidelines specific. *Bmj*, 328(7435), 343–345.
- Middleton, E. L., & Schwartz, M. F. (2012). Errorless learning in cognitive rehabilitation: a critical review. *Neuropsychological Rehabilitation*, 22(2), 138–168. doi:10.1080/09602011.2011.639619
- Morgan, E. E., Weber, E., Rooney, A. S., Grant, I., & Woods, S. P. (2012). Longer ongoing task delay intervals exacerbate prospective memory deficits in HIV-associated neurocognitive disorders (HAND). *Journal of Clinical and Experimental Neuropsychology*, 34(4), 416–427. doi:10.1080/13803395.2012.654764
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7(1), 105–125.
- Nelson, H. E., & Willison, J. R. (1991). *National Adult Reading Test (NART)*. Windsor: NFER-Nelson.
- Nowinski, J. L., & Dismukes, K. R. (2005). Effects of ongoing task context and target typicality on prospective memory performance: the importance of associative cueing. *Memory*, 13(6), 649–657. doi:10.1080/09658210444000313
- Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Yamadori, A., Frith, C. D., & Burgess, P. W. (2007). Differential involvement of regions of rostral prefrontal cortex (Brodmann area 10) in time- and event-based prospective memory. *International Journal of Psychophysiology*, 64(3), 233–46. doi:10.1016/j.ijpsycho.2006.09.009
- Ozgis, S., Rendell, P. G., & Henry, J. D. (2009). Spaced retrieval significantly improves prospective memory performance of cognitively impaired older adults. *Gerontology*, 55(2), 229–232. doi:10.1159/000163446
- Page, M., Wilson, B. A., Shiel, A., Carter, G., & Norris, D. (2006). What is the locus of the errorless-learning advantage? *Neuropsychologia*, 44(1), 90–100. doi:S0028-3932(05)00166-1

- Petrides, M. (1985). Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, 23(5), 601–614.
- Petrides, M. (1997). Visuo-motor conditional associative learning after frontal and temporal lesions in the human brain. *Neuropsychologia*, 35(7), 989–997. doi:10.1016/S0028-3932(97)00026-2
- Raskin, S. A. (2004). The Memory for Intentions Screening Test. *Journal of the International Neuropsychological Society*, 10(S1), 110.
- Roche, N. L., Fleming, J. M., & Shum, D. H. K. (2002). Self-awareness of prospective memory failure in adults with traumatic brain injury. *Brain Injury*, 16(11), 931–945.
- Roediger, H. L., 3rd, & Payne, D. G. (1982). Hypermnnesia: the role of repeated testing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 8(1), 66–72.
- Rous, R. S., Adams, M., Fish, J., Manly, T., & Adlam, A. (2012). Prospective Memory Intervention for Adolescents with Acquired Brain Injury: A Preliminary Study. Presented at the Ninth World Congress on Brain Injury, Edinburgh, UK.
- Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, 8(9), 657–661. doi:10.1038/nrn2213
- Schmitter-Edgecombe, M., & Wright, M. J. (2004). Event-based prospective memory following severe closed-head injury. *Neuropsychology*, 18(2), 353.
- Schönberger, M., & Ponsford, J. (2010). The factor structure of the Hospital Anxiety and Depression Scale in individuals with traumatic brain injury. *Psychiatry Research*, 179(3), 342–349. doi:10.1016/j.psychres.2009.07.003
- Scullin, M. K., McDaniel, M. A., & Einstein, G. O. (2010). Control of cost in prospective memory: evidence for spontaneous retrieval processes. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 36(1), 190–203. doi:10.1037/a0017732
- Shallice, T. (2003). Functional imaging and neuropsychology findings: how can they be linked? *NeuroImage*, 20, Supplement 1, S146–S154. doi:10.1016/j.neuroimage.2003.09.023
- Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, 114(2), 727–741.
- Shallice, T., & Burgess, P. W. (1996). The domain of supervisory processes and temporal organization of behaviour. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 351(1346), 1405–1411; discussion 1411–1412. doi:10.1098/rstb.1996.0124
- Shallice, T., & Burgess, P. W. (1997). *The Hayling and Brixton Tests*. Bury St Edmunds, Suffolk: Thames Valley Test Company.

- Shum, D., Fleming, J., Gill, H., Gullo, M., & Strong, J. (2011). A randomized controlled trial of prospective memory rehabilitation in adults with traumatic brain injury. *Journal of Rehabilitation Medicine, 43*(3), 216–223. doi:10.2340/16501977-0647
- Shum, D., Ungvari, G. S., Tang, W. K., & Leung, J. P. (2004). Performance of schizophrenia patients on time-, event-, and activity-based prospective memory tasks. *Schizophrenia Bulletin, 30*(4), 693.
- Shum, D., Valentine, M., & Cutmore, T. (1999). Performance of traumatic brain-injured individuals on time-, event-, and activity-based prospective memory tasks. *Journal of Clinical and Experimental Neuropsychology, 21*, 49–58.
- Smith, G., Del Sala, S., Logie, R. H., & Maylor, E. A. (2000). Prospective and retrospective memory in normal ageing and dementia: A questionnaire study. *Memory, 8*(5), 311–321. doi:10.1080/09658210050117735
- Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(3), 347.
- Smith, R. E. (2010). What Costs Do Reveal and Moving Beyond the Cost Debate: Reply to Einstein and McDaniel (in press). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(4), 1089–1095. doi:10.1037/a0019183
- Smith, S. J., Souchay, C., & Moulin, C. J. A. (2011). Metamemory and prospective memory in Parkinson's disease. *Neuropsychology, 25*(6), 734–740. doi:10.1037/a0025475
- Sohlberg, M. M., & Raskin, S. A. (1996). Principles of Generalization Applied to Attention and Memory Interventions. *The Journal of Head Trauma Rehabilitation, 11*(2), 65.
- Sohlberg, M. M., White, O., Evans, E., & Mateer, C. (1992a). Background and initial case studies into the effects of prospective memory training. *Brain Injury, 6*(2), 129.
- Sohlberg, M. M., White, O., Evans, E., & Mateer, C. (1992b). An investigation of the effects of prospective memory training. *Brain Injury, 6*(2), 139.
- Sopena, S., Dewar, B.-K., Nannery, R., Teasdale, T. W., & Wilson, B. A. (2007). The European Brain Injury Questionnaire (EBIQ) as a reliable outcome measure for use with people with brain injury. *Brain Injury, 21*(10), 1063–1068. doi:10.1080/02699050701630342
- Squire, L. R., Shimamura, A. P., & Graf, P. (1987). Strength and duration of priming effects in normal subjects and amnesic patients. *Neuropsychologia, 25*(1B), 195–210.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences, 93*(24), 13515–13522.
- Squires, E. J., Hunkin, N. M., & Parkin, A. J. (1997). Errorless learning of novel associations in amnesia. *Neuropsychologia, 35*(8), 1103–1111.

- Storey, K. (2010). Smart houses and smart technology: overview and implications for independent living and supported living services. *Intellectual and Developmental Disabilities, 48*(6), 464–469. doi:10.1352/1934-9556-48.6.464
- Svoboda, E., Richards, B., Leach, L., & Mertens, V. (2012). PDA and smartphone use by individuals with moderate-to-severe memory impairment: application of a theory-driven training programme. *Neuropsychological Rehabilitation, 22*(3), 408–427. doi:10.1080/09602011.2011.652498
- Sweeney, S., Kersel, D., Morris, R. G., Manly, T., & Evans, J. J. (2010). The sensitivity of a virtual reality task to planning and prospective memory impairments: group differences and the efficacy of periodic alerts on performance. *Neuropsychological Rehabilitation, 20*(2), 239–263. doi:10.1080/09602010903080531
- Tailby, R., & Haslam, C. (2003). An investigation of errorless learning in memory-impaired patients: improving the technique and clarifying theory. *Neuropsychologia, 41*(9), 1230–1240.
- Teasdale, T. W., Christensen, A. L., Willmes, K., Deloche, G., Braga, L., Stachowiak, F., ... Leclercq, M. (1997). Subjective experience in brain injured patients and their close relatives: A European Brain Injury Questionnaire study. *Brain Injury, 11*(8), 543–564.
- Terrace, H. S. (1963a). Discrimination learning with and without “errors.” *Journal of the Experimental Analysis of Behavior, 6*(1), 1–27. doi:10.1901/jeab.1963.6-1
- Terrace, H. S. (1963b). Errorless discrimination learning in the pigeon: effects of chlorpromazine and impiramine. *Science (New York, N.Y.), 140*(3564), 318–319.
- Terrace, H. S. (1963c). Errorless transfer of a discrimination across two continua. *Journal of the Experimental Analysis of Behavior, 6*(2), 223–232.
- Terrace, H. S. (2010). Defining the Stimulus - A Memoir. *Behavioural Processes, 83*(2), 139–153. doi:10.1016/j.beproc.2009.12.003
- Tulving, E., Kapur, S., Craik, F. I., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proceedings of the National Academy of Sciences of the United States of America, 91*(6), 2016–2020.
- Uttl, B., & Kibreab, M. (2011). Self-report measures of prospective memory are reliable but not valid. *Canadian Journal of Experimental Psychology, 65*(1), 57–68. doi:10.1037/a0022843
- Volle, E., Gonen-Yaacovi, G., de Lacy Costello, A., Gilbert, S. J., & Burgess, P. W. (2011). The role of rostral prefrontal cortex in prospective memory: A voxel-based lesion study. *Neuropsychologia, 49*(8), 2185–2198. doi:10.1016/j.neuropsychologia.2011.02.045
- Wechsler, D. (2001). *Wechsler Test of Adult Reading*. London: Pearson Assessment.

- Wechsler, D. (2011a). *Test of Premorbid Functioning - UK Version (TOPF UK)*. London: Pearson Assessment.
- Wechsler, D. (2011b). *Wechsler Abbreviated Scale of Intelligence - Second Edition (WASI-II)*. London: Pearson Assessment.
- Weiler, J., Suchan, B., Koch, B., Schwarz, M., & Daum, I. (2011). Differential impairment of remembering the past and imagining novel events after thalamic lesions. *Journal of Cognitive Neuroscience*, 23(10), 3037–3051. doi:10.1162/jocn.2011.21633
- Weinborn, M., Moyle, J., Bucks, R. S., Stritzke, W., Leighton, A., & Woods, S. P. (2013). Time-Based Prospective Memory Predicts Engagement in Risk Behaviors Among Substance Users: Results From Clinical and Nonclinical Samples. *Journal of the International Neuropsychological Society*, 19(03), 284–294. doi:10.1017/S1355617712001361
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, Texas: The Psychological Corporation.
- Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Ostensson, M.-L., Bartfai, A., & Klingberg, T. (2007). Computerized working memory training after stroke--a pilot study. *Brain Injury*, 21(1), 21–29. doi:10.1080/02699050601148726
- Wilson, B. A. (1991). Long-term prognosis of patients with severe memory disorders. *Neuropsychological Rehabilitation*, 1(2), 117–134.
- Wilson, B. A. (2009). *Memory Rehabilitation: Integrating Theory and Practice*. Guilford Press.
- Wilson, B. A., Baddeley, A. D., Evans, J. J., & Shiel, A. (1994). Errorless learning in the rehabilitation of memory impaired people. *Neuropsychological Rehabilitation*, 4, 307–326.
- Wilson, B. A., Cockburn, J., & Baddeley, A. D. (1985). *The Rivermead Behavioural Memory Test*. Thames Valley Test Company Reading.
- Wilson, B. A., Cockburn, J., Baddeley, A. D., & Hiorns, R. (1989). The development and validation of a test battery for detecting and monitoring everyday memory problems. *Journal of Clinical and Experimental Neuropsychology*, 11, 855–870.
- Wilson, B. A., Emslie, H., Foley, J., Shiel, A., Watson, P., Hawkins, K., ... Evans, J. J. (2005). *The Cambridge Prospective Memory Test*. London: Harcourt Assessment.
- Wilson, B. A., Emslie, H., Quirk, K., Evans, J. J., & Watson, P. (2005). A randomized control trial to evaluate a paging system for people with traumatic brain injury. *Brain Injury*, 19(11), 891–4. doi:10.1080/02699050500046571
- Wilson, B. A., Evans, J. J., Emslie, H., & Malinek, V. (1997). Evaluation of NeuroPage: a new memory aid. *Journal of Neurology, Neurosurgery & Psychiatry*, 63, 113–115.
- Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361–370.

- Zinke, K., Altgassen, M., Mackinlay, R. J., Rizzo, P., Drechsler, R., & Kliegel, M. (2010). Time-based prospective memory performance and time-monitoring in children with ADHD. *Child Neuropsychology*, 16(4), 338–349. doi:10.1080/09297041003631451
- Zogg, J. B., Woods, S. P., Weber, E., Iudicello, J. E., Dawson, M. S., & Grant, I. (2010). HIV-associated Prospective Memory Impairment in the Laboratory Predicts Failures on a Semi-naturalistic Measure of Health Care Compliance. *The Clinical Neuropsychologist*, 24(6), 945–962. doi:10.1080/13854046.2010.501343

7. Appendices

7.1. Procedures for Determining Mental Capacity to Consent

The following procedures were approved by the NHS Research Ethics Committee:

Informed consent will be obtained by the Chief Investigator in Session 1, and confirmed at the beginning of Session 2. This will involve the explanation of the purpose of the research and procedures involved, and all other relevant information as set out in the Participant Information Sheet. Though all participants are initially assumed to have capacity to provide informed consent, due to the nature of the population being studied (i.e. people with memory problems), it may transpire that some participants lack this capacity. The Department of Health guidance states that if a research study is related to the 'impairing condition' causing the incapacity, and if excluding those unable to consent for themselves would reduce the effectiveness of the study, then such participants should be included. As memory problems are central to the aims of the study, and to exclude people lacking capacity would likely introduce a sampling bias towards more mildly affected individuals, this project should not exclude participants who lack capacity to provide informed consent. If a potential participant's capacity to provide informed consent is in question (in either Session 1 or 2), then a Capacity Assessment will be completed. The Capacity Assessment will include consideration of the following criteria:

- Does the potential participant understand the relevant information?
- Can s/he retain that information?
- Can s/he use/weigh that information in the process of making a decision to participate?
- Can s/he communicate that decision reliably?

In practice, the assessment will be conducted over an initial telephone call and, contingent on the person's verbal consent, the research appointments. The purpose of the study and demands of participation will be explained to potential participants when they are contacted by telephone to receive an introduction to the study and arrange an initial appointment. If the participant provides initial verbal consent to attend an assessment session, then the project will proceed, otherwise the process will be terminated at this point. At the initial session, the formal procedures regarding obtaining informed consent will be initiated. Participants will be asked to state their understanding of the project and what they will be asked to do at the outset of the session in the course of the introductory conversation, which will allow for examination of the first two criteria. The participant will be asked to state what any possible advantages or disadvantages may be, to address the third criteria, and then to reach a decision regarding whether or not he/she wishes to take part, to address the final criteria. If all criteria

are satisfied, then the research procedures will be initiated. If there is a doubt regarding any criteria, the potential participant's Personal Consultee (e.g. next of kin or other person who knows the potential participant well and is not acting in a professional capacity) will be contacted and provided with information regarding the study, and asked to make an informed decision to consent on behalf of the potential participant, in his/her best interests. If no Personal Consultee is available and the potential participant is keen to participate, we will consider using a Nominated Consultee (i.e. someone independent of the research team who knows the person in a professional context). If consent is obtained, the assessment will commence.

As participants will be recruited on the basis of having memory problems, it is important to thoroughly address the second criteria, regarding retention of information. Within this assessment, retention is addressed by means of the delay between the provision of initial information over the telephone and enquiry regarding the potential participant's understanding of the purpose of the appointment at the first research session. If the participant can remember what they research project involves, then it would be clear evidence of retention of information. If a participant cannot remember the purpose of the appointment or aims of the research then it may indicate a problem at the level of the 'retention' criteria, and the participant's Personal Consultee will be contacted. The consent procedure will be repeated at the beginning of the second research session to address the following situations:

- a participant who had previously provided informed consent has forgotten information relevant to the decision to participate (i.e. it later became apparent that the participant lacked capacity to make the decision to participate)
- a participant previously lacking capacity for this decision has regained that capacity
- a participant has changed their mind about the decision to participate.

An equivalent process will take place for the supplementary small-n study, though it will be more detailed to reflect the greater time commitment and increased demands related to this study.

Assessments will be undertaken by Dr Fish, under the supervision of Professors Morris and Kopelman. Dr Fish has training and experience in conducting such assessments as part of her Doctorate in Clinical Psychology, and experience in obtaining informed consent for participation in research in the course of her research career. These procedures are in accordance with the Mental Capacity Act (2005), the British Psychological Society Short Reference Guide for Psychologists on the MCA, and the Department of Health's Guidance on the inclusion of people who lack capacity to provide informed consent in research.

7.2. Background Information on Study Participants

Table 1.7. Table showing demographic, medical, neuropsychological and questionnaire data for each participant, and group summaries where relevant.

ID	Sex	Age	Yrs Educ	Age at Onset	Aetiology	Scan report summary	IQ Z-score	Memory Z-score	IQ-memory discrep.	PRMQ		HADS		EBIQ (vs Brain Injury Group)		EBIQ (vs Control Group)	
										Pro	Retro	Anx	Dep	Cog	Core	Cog	Core
1	M	49	13	24	Head injury/alcohol	Steno-occlusive disease, M1 MCA	0	-1.67	-1.67	<-3.0	-1.50	-0.6	-0.7	-0.50	-0.24	-1.66	-1.05
2	M	52	11	28	Diabetes	Prominent ventricles and subarachnoid spaces, posterior fossa volume loss	1	-1.58	-3	-2.25	-2.67	-2.3	-2.3	-1.03	-2.20	-2.40	-3.40
3	M	69	12	65	SVD	Mild-moderate Small Vessel Disease	0.25	-1.67	-1.58	+1.20	0.9	-2.3	-0.2	0.76	0.83	0.08	0.23
4	M	60	11	57	Stroke	Left PCA stroke, volume loss L fusiform and peri-hippocampal gyri and posterior hippocampus	-0.25	-1.67	-2	-0.4	-0.33	-0.6	-0	0.04	0.50	-0.91	-0.17
5	M	61	9	48	Stroke and SVD	L thalamic and cerebellar atrophy	-0.33	-3	-1.67	-0.60	0.100	0	-0.7	-1.03	0.34	-2.40	-0.36
6	M	45	17	44	Hypoxia	Multiple infarcts, L frontal, temporal and parietal	1.67	-2.33	-4.34	-0.60	+0.55	-0.1	-0	0.22	0.75	-0.66	0.13
7	M	55	11	46	Hypoxia	Not available	0.55	-2.33	-2.55	<-3.0	-2.67	-2.3	-2.3	-2.46	-2.93	-4.38	-4.28
8	F	65	16	52	TLE	High signal over left superior and middle temporal gyri	1.67	-1.33	-2.67	-0.40	-0.5	0	-0.2	-0.32	0.25	-1.41	-0.46
9	M	51	11	48	Stroke	Bilateral inferior cerebellar infarcts	-0.75	-3	-1.58	-2.75	-0.9	-1.3	-1.1	-0.32	0.17	-1.41	-0.56
10	M	52	11	40	Epilepsy/stroke	Bilateral hippocampal sclerosis	-0.5	-3	-1.5	+0.60	-3.00	-2.3	-0.7	-2.11	-1.95	-3.89	-3.11
11	M	59	11	55	Stroke	R parietal infarct, R hippocampal atrophy, frontal grey matter volume loss	0.75	-3	-2.75	-2.75	-2.50	-0.9	-2.3	-2.64	-2.28	-4.63	-3.50
12	M	47	11	40	Hypoxic	Cerebellar atrophy, hippocampal atrophy	-0.75	-3	-1.58	+1.00	0.1	+1	-0.2	0.58	1.40	-0.17	0.91
13	M	52	10	50	Stroke	Two infarcts, affecting L occipital and medial posterior temporal lobes bilaterally, and R thalamus	0.67	-1.33	-3.34	<-3.0	-1.9	-1.6	-1.9	-1.75	-1.71	-3.39	-2.81
14	F	38	12	35	Temporal lobe epilepsy	Not available	-0.25	-2.67	-2.42	-1.25	-0.1	+1	-0.6	0.04	-0.15	-0.91	-0.95
Group M							0.27	-2.26	-2.05	-0.78	-1.73	-0.88	-0.94	-0.75	-0.52	-2.01	-1.38
Group SD							0.81	0.69	1.43	1.47	1.37	1.18	0.89	1.12	1.40	1.55	1.67

Abbreviations: Anx, Anxiety; Cog, Cognitive; Dep, Depression; EBIQ, European Brain Injury Questionnaire; HADS, Hospital Anxiety and Depression Scale; MCA, Middle Cerebral Artery; PCA, Posterior Cerebral Artery; PRMQ, Prospective and Retrospective Memory Questionnaire; Pro, Prospective; R, right; Retro, Retrospective SVD Small Vessel Disease; Yrs Educ, Years of Education;

Note that age, years of education, and age at onset are given in years, whereas all other scores are presented as z-scores for ease of reading. Z-scores were computed from the relevant test norms, and for the questionnaires, on the basis of normative data provided in Crawford et al. (2003) for the PRMQ, Crawford et al. (2001) for the HADS, and Teasdale et al. (1997) for the EBIQ. Scores are scaled such that negative values represent impairment relative to the population mean.

Service-Related Research Project

An Audit of Referrals Accepted by a Tier 4 CAMHS Team Pre- and Post- Publication of NICE Guidance for Conduct Problems.

Supervised by Dr Anouk Houdijk, Clinical Psychologist, National & Specialist CAMHS Conduct,
Adoption and Fostering Team, South London & Maudsley NHS Foundation Trust.

Table of Contents

List of Tables	97
List of Figures	98
List of Abbreviations	99
1. Abstract.....	100
2. Introduction	101
2.1. Conduct Disorders: Characteristics, Causes, Prevalence and Outcomes.....	101
2.2. Parent Education & Training Programmes for Managing Conduct Disorders	102
2.2.1. Guidance on Best Clinical Practice	102
2.2.2. Parent Training in the Context of Adoption and Fostering.....	103
2.2.3. Parent Management Training in a National & Specialist CAMHS Setting.....	104
2.3. Aims of the Audit	105
3. Method	106
3.1. Participants	106
3.2. Data Collection.....	106
3.3. Measures.....	106
3.3.1. Demographic and Other Background Information	106
3.3.2. Referral Information	106
3.3.3. Presenting Problems	107
3.3.4. Psychometric Assessment Results	107
3.3.5. Diagnostic Information	107
3.3.6. Strengths and Difficulties Questionnaire Scores.....	107
3.3.7. Treatment Plan	108
3.3.8. Complexity Index.....	108
3.4. Statistical Analysis.....	108
4. Results.....	110
4.1. Demographic and Related Characteristics of the Sample	110
4.1.1. Referral Source.....	110

4.1.2.	Gender	110
4.1.3.	Ethnicity	111
4.1.4.	Age at Referral	113
4.1.5.	Parental Mental Health Problems and Learning Disability	114
4.1.6.	Comparison of Child Factors over the Three Time Periods	117
4.2.	Examination of Complexity Index by Time Period	118
4.3.	Examination of Treatment Plans by Time Period.....	119
4.3.1.	Number of Treatments	119
4.3.2.	Examination of Treatment Plans by Time Period.....	120
4.3.2.1.	Parent-Child Game.....	121
4.3.2.2.	Anger Management.....	121
4.3.2.3.	Behavioural Family Work.....	121
4.3.2.4.	Medication.....	121
4.3.2.5.	School Liaison.....	122
5.	Discussion.....	123
5.1.	Summary of Results and their Relation to the Hypotheses	123
5.2.	Complexity and the TA: Implications for Families and Services	125
5.3.	The Impact of the TA on the A&FS.....	126
5.4.	Suggestions for Further Investigation.....	126
5.5.	Limitations of the Current Project	127
6.	References	129
7.	Appendix 1: Data Collection Form	132
8.	Appendix 2: Supplementary Statistical Analyses	134
9.	Appendix 3: Supplementary Information on Primary Analyses.....	140

List of Tables

Table 2.1. NICE (2006) summary of outcomes from parent-training programs in children with conduct disorders.	102
Table 2.2. The number of referrals accepted by the CPS and A&FS in each time period.....	106
Table 2.3. The ethnicity of referrals to each clinic over each time period.	111
Table 2.4. The mean age (in years) of referrals to each service in each time period, with standard deviations in parentheses.....	114
Table 2.5. The occurrence of any parental mental health problem or learning disability, according to clinic and time period.	114
Table 2.6. Diagnostic frequencies within the A&FS over the three time period.	118
Table 2.7. Percentage of cases with PCG as part of the treatment plan over the three time periods and two clinics	121
Table 2.8. Percentage of cases with Anger Management as part of the treatment plan by time period and clinic.....	121
Table 2.9. Percentage of cases with Behavioural Family Work as part of the treatment plan by time period and clinic.	121
Table 2.10. Percentage of cases with Medication as part of the treatment plan by time period and clinic.	122
Table 2.11. Percentage of cases with Medication as part of the treatment plan over the three time periods and two clinics.	122
Table 2.12. Percentage of cases with a diagnosis at referral, and a diagnosis after assessment, by time period and clinic.....	134
Table 2.13. A summary of analyses conducted on referral diagnoses by time period and clinic.	134
Table 2.14. A summary of the analyses conducted on the outcome diagnoses for each clinic over time.....	135
Table 2.15. A summary of analyses conducted on referral problems over by time period and clinic.	135
Table 2.16. A summary of analyses conducted on presenting problems over three time periods per clinic.....	135
Table 2.17. Descriptive statistics for SDQs by Service and Time Period.....	137
Table 2.18. Descriptive Statistics of Psychometric Assessment Scores by Service and Time Period.....	139
Table 2.19. Descriptive statistics of the complexity index by Service and Time Period.	140
Table 2.20. Number of treatments planned for each service in each time period.	140

List of Figures

Figure 2.1. Graph showing the number of referrals from CAMHS versus non-CAMHS sources in both the CPS and A&FS over the three time points.....	110
Figure 2.2. Distribution of Ethnicity by Time Period within the Conduct Problems Clinic (Panel A) and the Adoption and Fostering Clinic (Panel B).....	112
Figure 2.3. Chart showing the age at referrals of cases within each service at each time period.	113
Figure 2.4. Pie charts showing the change in parental mental health problems in the CPS over the three time periods.	116
Figure 2.5. Composite Complexity Index Scores according to Service and Time Period.....	119
Figure 2.6. The number of treatments planned per family by service and time period.	120

List of Abbreviations

A&FS	Adoption and Fostering Service
ADHD	Attention Deficit Hyperactivity Disorder
ANOVA	Analysis of Variance
ASD	Autism Spectrum Disorder
CAMHS	Child and Adolescent Mental Health Services
CAFT	Conduct, Adoption and Fostering Team
CBT	Cognitive Behavioural Therapy
CD	Conduct Disorder
CPS	Conduct Problems Service
ICD-10	International Classification of Diseases - Tenth Edition
NICE	National Institute for Health and Clinical Excellence
ODD	Oppositional Defiant Disorder
ONS	Office for National Statistics
PCG	Parent-Child Game
RCT	Randomised Controlled Trial
SDQ	Strengths and Difficulties Questionnaire
SLT	Social Learning Theory
TA	Technology Appraisal
WIAT-II	Wechsler Individual Attainment Test – Second UK Edition

1. Abstract

In 2006 the National Institute of Health and Clinical Excellence (NICE) published a Technology Appraisal (TA) recommending group parenting interventions as the first-line treatment for conduct problems in children aged under 12 years. Individual-format interventions were recommended for treatment-refractory, complex, or difficult to engage cases only. This should have resulted in increased availability of parent training programs in Child and Adolescent Mental Health Services (CAMHS), particularly at Tiers 2-3, with more complex cases seen at Tier 4. This audit retrospectively reviewed case notes to investigate the impact of the TA on referrals to a Tier 4 Conduct Problems Service (CPS) over three two-year time periods: two years prior to publication of the TA, 0-2 years after the TA, and 2-4 years after the TA. Cases referred to the associated Adoption & Fostering Service (A&FS) formed a comparison group. The main finding was that the overall complexity of referrals accepted by the CPS increased following publication of the TA, whereas referrals accepted by the A&FS remained stable. Further analyses concerning potential sources of complexity identified that there had been significant increases in the proportion of CPS referrals from CAMHS (suggesting engagement problems or treatment non-response), and in the proportion of cases with maternal mental health problems, compared with referrals to the A&FS. Though an increase in comorbid disorders was also expected, this was not observed. The impact of the change on treatment plans indicated an increase in the proportion of cases receiving specialist parenting interventions. The results of the audit demonstrate that TA has had an impact on the CPS caseload, and this suggests that the recommendations are being followed locally. This increase in complexity has implications for the service in terms of necessary resources, innovation in clinical practice, and in monitoring outcomes.

2. Introduction

2.1. Conduct Disorders: Characteristics, Causes, Prevalence and Outcomes

Conduct disorders are characterised by a pattern of dissocial, aggressive or defiant behaviour in childhood. For a diagnosis to be made within ICD-10, at least four discrete behaviours that markedly deviate from age-related expectations must have been present for at least six months. Such behaviours might include fighting, bullying, cruelty to people or animals, destructiveness to property, fire-setting, stealing, lying, truancy, temper tantrums, defiant behaviour and persistent disobedience. If the child's behaviours do not involve violation of the rights of others, then a 'lesser' diagnosis of Oppositional Defiant Disorder (ODD) may be diagnosed. Other subtypes of Conduct Disorder (CD) include that confined to the family context (i.e. behaviours are not apparent at school or with other caretakers), and unsocialised CD (i.e. where appropriate peer relations are absent).

It is well established that both biological and psychosocial factors are involved in the development of CDs. Scott (1998) reported that the primary causal factor is exposure to a harsh and inconsistent style of parenting, with contributions from secondary influences such as child hyperactivity and reduced IQ (which is at a group level one standard deviation below the population mean). Moffitt (2005) examined the large research literature on twin studies of antisocial behaviour, and concluded that genetic influences account for 40-50% of the observed variance in antisocial behaviour, environmental influences 15-20%, and the remainder being explained by 'non-shared experiences'. However, there are also important gene-environment interactions that impact upon the development of CDs. Caspi et al (2002) found that a particular low-activity polymorphism of the Monoamine Oxidase A gene makes CD more likely to occur in maltreated children compared with both maltreated children without that polymorphism and non-maltreated children with the same polymorphism. Additionally, Bohman (1996) reported results from an adoption study, whereby better parenting almost completely removes the negative impact of a genetic risk for criminality. This is a striking demonstration of the influence exerted by psychological factors on the development of CD.

Conduct problems during childhood are associated with a wide variety of negative outcomes later in life, including poor educational outcomes, unemployment, low occupational status, marital and social adjustment problems, an increased likelihood of hospitalisation for both physical and mental health problems, and criminality (Carr, 2006). It has been estimated that preventing CD saves £150,000 per case in lifetime costs, around two thirds of which are savings associated with prevention of crime and imprisonment (Friedli & Parsonage, 2007).

Early identification of, and effective intervention for, conduct disorders are therefore of great importance within the educational, economic, and health care systems.

The Office of National Statistics states that CD has a prevalence rate of 6% in the UK population, whereas prevalence rates for ODD vary from 4-14%. Between 2-4 times as many males as females are affected. Conduct problems are particularly prevalent in fostered and adopted children, with estimates varying from 20% of adopted children in the US (Simmel, Brooks, Barth, & Hinshaw, 2001), and as high as 37% of children in foster placements in the UK (Tapsfield & Collier, 2005). Taking into account the current social context of adoption and fostering, whereby the vast majority of children removed from their birth families have experienced significant early neglect and/or other abuse, this elevated prevalence is not surprising. Additionally, there is a tradition for behavioural problems in looked after or adopted children to be conceptualised within an attachment theory framework, rather than through social learning theory, as would be commonplace for children living with their birth families (Nilsen, 2003). As a result, CDs may be under-diagnosed in this population (NICE, 2010), and the actual prevalence of CDs in children in adoptive or foster placements may be much higher. The prevalence of conduct problems, and the adversity associated with them, makes the need for effective interventions extremely clear.

2.2. Parent Education & Training Programmes for Managing Conduct Disorders

2.2.1. Guidance on Best Clinical Practice

On 21st July 2006 the National Institute for Health and Clinical Excellence (NICE) published a Technology Appraisal (TA) recommending group-based parent-training/education programmes based upon principles of Social Learning Theory (SLT; Bandura, 1976) as a first-line treatment in the management of conduct problems in children aged 12 years or younger. This was based upon a substantial evidence base consisting of 16 published systematic reviews of parenting programmes, with 6 of those reviews being considered of high quality, and all six reporting positive outcomes. Statistically significant positive effects for long-term outcomes were also reported in 11/15 rigorous studies. The results from 25 Randomised Controlled Trials (RCTs) were also considered, including 19 of parent training versus an active rather than passive control condition, with 50% showing a significant beneficial effect of parenting programmes and the remaining 50% reporting neutral. An updated meta-analysis (Dretzke et al., 2009) computed the effect size from 57 RCTs as .67 (95% CI .42-.91) for parent-rated outcomes and .44 (95% CI: .23-.66) for independently rated outcomes.

Table 2.1. NICE (2006) summary of outcomes from parent-training programs in children with conduct disorders.

Guidance applies to children aged up to 12 years. Comparison groups include psychotherapy, CBT, social skills training, play therapy, music therapy, art therapy and occupational therapy.

Outcomes from Parent Training programs have included:

Child:

- Improved behaviour (parent and teacher report, independent observation)
- Improved self-esteem
- Reduced rates of school exclusion
- Reduced rates of offending

Parent:

- Improved self-esteem
- Reduced parental stress, anxiety and depression
- Reduced criticism towards child

Family

- Change in attitudes
- Improved family adjustment
- Increased placement stability

The TA states that individual parenting programmes are necessary for families with complex needs or who it has been difficult to engage in group programmes. Factors associated with non-completion of programmes include young age, low socio-economic status, less social support, higher life stress, less education, higher maternal depression on the Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995), and higher parental dysfunction, and hence the presence of such factors may indicate that an individual format is warranted (NICE, 2006).

2.2.2. Parent Training in the Context of Adoption and Fostering

The TA does not specifically refer to children in adoptive/foster placements (though note that the forthcoming NICE Guidance on Conduct Disorders will include this population; NICE, 2010). Although there is evidence that parent management training is effective in helping foster carers to manage conduct problems in the children they look after (Bywater et al., 2010; Pallett, Scott, Blackeby, Yule, & Weissman, 2002; Rushton, Monck, Leese, McCrone, & Sharac, 2010), this evidence comes from smaller and less methodologically rigorous studies than those included in the TA review. There may be a number of practical reasons why the evidence-base for effective interventions for fostered or adopted children lags behind that for children in their birth families. For instance, there are fewer adopted or fostered children and hence fewer relevant services and less research funding allocated to the topic. However, a further contributing factor may be the aforementioned tradition for behavioural problems in looked after or adopted children to be conceptualised within an attachment theory framework. Despite the large body of empirical research on attachment (initiated by e.g. Ainsworth, 1979; Bowlby, 1953), therapies focussed on attachment, such as Theraplay (Booth & Jernberg, 2010) and Dyadic Developmental Psychotherapy (Hughes, 2004), seem to have a weaker association with the scientific tradition than behavioural and cognitive-behavioural approaches. Nilsen (2003) stated, with reference to the United States, that in adoption

services attachment is ill-defined and used as a catch-all term to describe and explain any aspect of a child's behaviour, and it bears little resemblance to attachment theory in its empirical form. Further the lack of research demonstrating the efficacy of attachment-based therapies is striking, with a preponderance of case reports of esoteric approaches (e.g. using "regression" techniques such as bottle-feeding a child who is well able to feed him/herself; Hughes, 2004), and the one UK-based RCT showing no benefit of additional training on attachment and communication when compared with training-as-usual for foster carers (Minnis, Pelosi, Knapp, & Dunn, 2001). In contrast, effective behavioural interventions such as Multidimensional Treatment Foster Care have been found to impact positively on attachment behaviours (Fisher & Kim, 2007). As there can be a failure to recognise and diagnose Conduct Disorders in adopted or fostered children (NICE, 2010), it is also likely that such children are denied access to potentially highly effective treatments.

2.2.3. Parent Management Training in a National & Specialist CAMHS Setting

The Conduct, Adoption and Fostering Team (CAFT), is part of the National and Specialist (Tier 4) Child and Adolescent Mental Health Service within South London and Maudsley NHS Foundation Trust. CAFT comprises two linked clinics, the Conduct Problems Service (CPS) and the Adoption and Fostering Service (A&FS), each offering specialist assessment, formulation and treatment. The link between the clinics is through the child's likely exposure to parenting practices that are either suboptimal for, or mismatched to, their individual needs. The team specialises in Parent Management Training for conduct problems, based on the "Parent Child Game" (PCG) program described by McMahon and Forehand (2005). PCG is run on a single-family basis, incorporating live coaching from clinicians, and as such constitutes a specialist programme appropriate for families who have not engaged with group programmes or who have more complex needs, including adoption or fostering as a source of complexity. Group-format approaches are more usually found in community and social services, along with Tier 3 Child and Adolescent Mental Health Services (CAMHS) and certain voluntary agencies.

As the evidence for parenting programmes in the management of conduct problems is so strong and the NICE guidance so clear, group-based interventions should be increasingly available in local CAMHS and other relevant community and voluntary agencies, with referrals to specialist services such as CAFT being reserved for cases who have not engaged with prior parenting work or who have complex needs. A previous study completed as part of a Doctorate in Clinical Psychology (Baldock, 2008) examined this issue in the two years prior to and following publication of the TA. A significant increase in overall complexity of cases seen after the TA was found, when complexity was indexed on an ordinal scale comprising the child

factors of intellectual ability (any index score <80), specific learning disability, comorbid disorders, and pre-natal or perinatal complications, along with parental mental health problems. However, no significant differences were observed in terms of either treatment history (specifically parenting work), or referral source (where an increase in referrals from CAMHS and decrease in referrals from GPs was expected). The impression within the team prior to the commencement of the current audit was that in the intervening time-period, referrals indeed concerned more complex cases, and this had implications for service delivery and resourcing.

2.3. Aims of the Audit

This audit aimed to explore changes in the characteristics of cases referred, and subsequent treatment plans, according to the following hypotheses generated through team discussion:

- The post-TA time periods will include more cases with maternal mental health problems and a history of treatment non-response than the earlier time period.
- The post-TA time periods will include more cases with comorbid disorders, particularly Attention Deficit Hyperactivity Disorder (ADHD) and Autistic Spectrum Disorders (ASD), than pre-TA.
- The post-TA time periods will include cases of increased overall complexity in comparison with the earlier time period, when complexity is indexed by the presence of maternal mental health problems, comorbid disorders, previous contact with CAMHS, IQ < 80, and literacy problems.
- The post-TA time periods will include more complex planned interventions, specifically that individual families will be offered a greater number and variety of treatments, compared pre-TA.

These hypotheses were tested in both CPS and A&FS referrals over three two-year time periods: Two years pre-TA (1st August 2004 – 31st July 2006), 0-2 years post-TA (1st August 2006 – 31st July 2008), and 2-4 years post-TA (1st August 2008 – 31st July 2010).

3. Method

3.1. Participants

Participants in this project were 264 children who had been referred to the Conduct Problems Clinic or the Adoption and Fostering Clinic and whose referral had been accepted. They were categorised according to the Service to which they had been referred, and the time period in which they had been referred.

Table 2.2. The number of referrals accepted by the CPS and A&FS in each time period.

Service	Time Period			Total
	2 years pre	0-2 years post	2-4 years post	
CPS	33	33	39	105
A&FS	46	50	35	131

3.2. Data Collection

As the project used retrospective case-review methodology, it was not necessary to obtain consent to use data from parents or guardians. The project was approved by South London and Maudsley CAMHS Clinical Audit Committee. Data relating to referrals in the first and second time periods had been collected as part of a previous project (Baldock, 2008). Data for the third time period were collected from electronic records using an amended version of the proforma from the original study.

3.3. Measures

Information was collected from case notes in the following categories (see Appendix 1 for the data collection form, though note that not all variables were analysed, some e.g. placement-related information were included for the purposes of future audits).

3.3.1. Demographic and Other Background Information

Information on date of birth, age, sex, and ethnicity were recorded, with ethnicity coded according to Office of National Statistics (ONS) categories. Parental mental health problems and learning disabilities were also recorded (for birth and adoptive/foster parents though the latter were not used here).

3.3.2. Referral Information

Information gathered on the referral included the referral type (Conduct or Adoption & Fostering), referral source (Tier and geographical location), any referral diagnosis, and the referral problems according to the following categories: conduct symptoms, attachment problems, ADHD symptoms, ASD symptoms, mood disorder, anxiety disorder, enuresis or encopresis, learning disability, other.

3.3.3. Presenting Problems

Information from the intake report was coded to summarise the presenting problems for each case, using the same categories as above.

3.3.4. Psychometric Assessment Results

Standardised scores from assessment of intellectual functioning and academic attainment were recorded. The most frequently used IQ tests were the Wechsler Abbreviated Scale of Intelligence for the majority of referrals of children aged 6 years or over, a prorated version of the Wechsler Preschool and Primary Scale of Intelligence 3rd edition for children aged 3-5 years, and in some cases the Wechsler Intelligence Scale for Children, 4th Edition where for example a more detailed assessment of cognitive strengths and weaknesses was indicated. For the current purposes the following summary indices were recorded: Full Scale IQ, Verbal IQ (or Verbal Comprehension Index), and Performance IQ (or Perceptual Reasoning Index). The measures of scholastic attainment were usually the Word Reading and Spelling subtests of the Wechsler Individual Achievement Test, Second UK Edition (WIAT-II) being word reading and spelling from the WIAT-II. Standard or index scores were used in the analysis.

3.3.5. Diagnostic Information

Post-assessment diagnoses were coded according the most frequently used diagnostic categories, along with a “total number of diagnoses” variable.

3.3.6. Strengths and Difficulties Questionnaire Scores

Where available, the Strengths and Difficulties Questionnaire (R. Goodman, 1997) scores were also recorded. The SDQ is a 25-item measure consisting of 5 subscales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behaviour. Scores from all subscales except prosocial behaviour can be summed to give a total difficulties score. There is a supplementary ‘impact’ subscale consisting of 8 items related to the severity, chronicity, distress and burden associated with the child’s difficulties. Separate versions are available for parent and teacher ratings of children aged 3-16 (the 3-4 year old version being slightly different from the version for 4-16 year olds), and another version for self-ratings of children aged 11-17 years. Excellent UK normative data is available and the questionnaires and related scoring information are available free online. The SDQ is a very widely-used instrument, it has adequate reliability (internal consistency, cross-rater, and test-retest reliabilities), and scores over the 90th percentile are strongly associated with independently diagnosed psychiatric disorders, which is evidence of its validity (A. Goodman & Goodman, 2009).

3.3.7. Treatment Plan

Information on the treatment plan was coded according to the most frequent interventions used (e.g. PCG, CBT, life story work), and work with wider systems such as school and Social Services liaison. Finally, a “total number of treatments” variable was created for each case.

3.3.8. Complexity Index

There are numerous potential definitions of the term “complexity” in relation to clinical cases, and the NICE guidance does not offer any definition, it only lists the aforementioned factors associated with programme non-completion. Yates, Garralda and Higginson (1999) developed a measure of case complexity for use within CAMHS (the Paddington Complexity Scale), which included the type, severity and chronicity of the condition, presence of co-morbid psychiatric and physical health problems, involvement of external agencies, type of school attended, family structure and attitude to services, prior contact with services, and presence of child protection concerns. However, the psychometric properties of this measure are not established, and data on many of the composite variables was not available in the existing two time-points of the dataset. Therefore, the decision was made to adapt the complexity index used by Baldock (2008), by assigning a score of 1 for the presence of each of the following: IQ<80, specific learning disability, presence of other co-morbid disorder diagnosed after assessment, any parent with a mental health problem or learning disability, CAMHS referral¹⁸. The resulting score consequently ranges from 0-5, with higher values indicating greater complexity. The index could be computed for 96.5% of the sample (missing data included 12 cases in time periods 1 and 2 for whom no diagnostic data were available).¹⁹

3.4. Statistical Analysis

For the categorical variables, contingency tables were constructed of the 3 time periods by n levels of the category in question, for the CPS and A&FS separately. To determine whether or not the distribution of each categorical variable varied over the three time periods, chi square statistics were obtained. In analyses where the expected values of more than 20% of the cells were less than 5, Fisher’s exact tests were completed. Where statistically significant differences were observed on these ‘omnibus’ tests, frequency tables were inspected to determine the qualitative nature of the change (Robson, 1994). Where such visual inspection did not sufficiently clarify the nature of the difference, the original comparisons were followed

¹⁸ This is a slight variation on the index used in the previous audit, as data on pre- or peri-natal complications was not available. This variable had been included on the grounds of its association with neurodevelopmental disorders, but here any such disorder would be accounted for in the ‘number of diagnoses’ variable.

¹⁹ A 4-component complexity index could be computed for 100% of the sample but as the results did not differ from the 5-component analysis it is not described further.

up with finer-grained analyses (i.e. of smaller contingency tables, an approach that is analogous to post-hoc comparisons after a significant result from an analysis of variance, ANOVA). Where such analyses were conducted, alpha values were corrected for multiple comparisons using Sidak's adjustment, to minimise the possibility of making Type II errors.

Boxplots of the continuous variables were examined for normality of distribution. As all variables were approximately normally distributed, they were analysed using analysis of variance (ANOVA). All analyses were conducted using PASW Statistics release 18.0 (SPSS Inc, 2009).

4. Results

4.1. Demographic and Related Characteristics of the Sample

4.1.1. Referral Source

The number of referrals to the CPS from CAMHS (Tiers 2-4) as opposed to non-CAMHS sources (primary care, social services and 'other') almost doubled in Time Period 3 compared with Time Periods 1 and 2. This change was statistically significant within the CPS ($\chi^2=7.737$, $p=.021$, $df=2$), but not the A&FS ($\chi^2=2.353$, $p=.308$, $df=2$), see Figure 2.1.

Although referral source is not a perfect indicator of previous treatment failure, it is likely that cases referred from other mental health services have either failed to engage with or otherwise respond to this treatment within those services, and as such, the pattern fits with the prediction that referrals after the TA would concern more complex cases for whom group-format training had been unsuccessful.

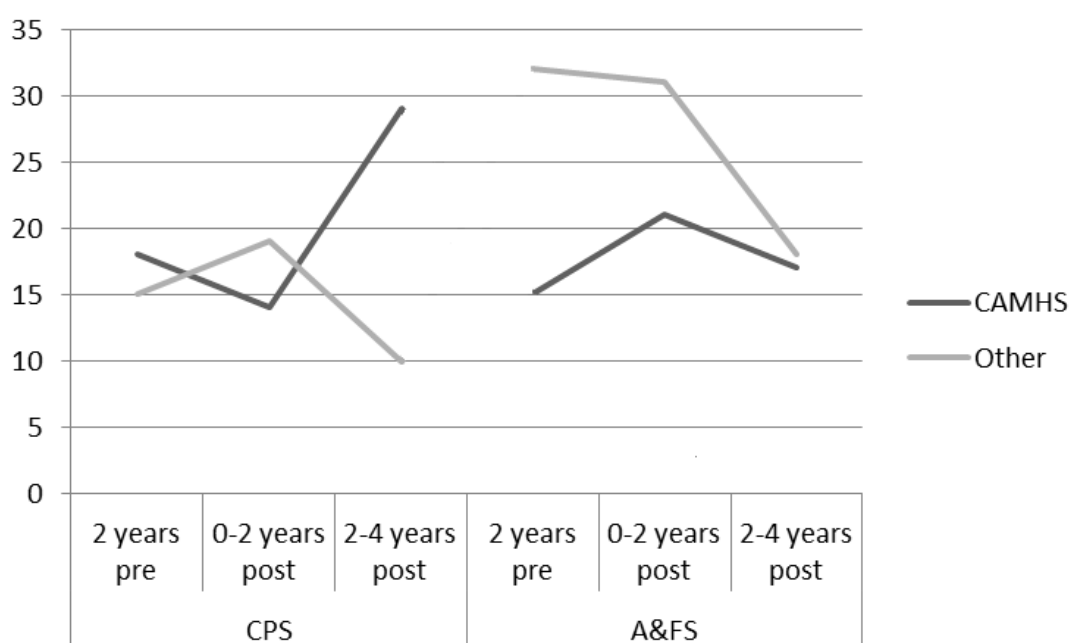


Figure 2.1. Graph showing the number of referrals from CAMHS versus non-CAMHS sources in both the CPS and A&FS over the three time points.

4.1.2. Gender

As would be expected in a clinic focussing on externalising disorders, many more males than females were referred (CPS 86% males; A&FS 62% males). However, there were no significant differences in the gender distribution over the three time periods in either the Conduct Clinic (Fisher's exact=.642, $p=.828$, $df=2$), or the A&FS (Fisher's exact=3.268, $p=.194$, $df=2$).

4.1.3. Ethnicity

There were significant differences in the distribution of the ethnicity of referrals over the time periods were found in both the CPS (Fisher's exact=22.345, $p=.004$, $df=10$), and the A&FS (Fisher's exact=29.668, $p<.001$, $df=10$).

Table 2.3. The ethnicity of referrals to each clinic over each time period.

Service	Ethnic Category	Time Period			Total
		2 years pre	0-2 years post	2-4 years post	
CPS	White	15	16	14	45
	Mixed	2	5	8	15
	Asian / Asian British	0	1	0	1
	Black / Black British	4	5	15	24
	Other Ethnic Group	1	1	1	3
	Not stated/missing	11	5	1	17
	Total	33	33	39	105
A&FS	White	14	35	21	70
	Mixed	8	7	2	17
	Asian / Asian British	1	0	2	3
	Black / Black British	3	2	4	9
	Other Ethnic Group	1	0	3	4
	Not stated/missing	20	8	3	31
	Total	47	52	35	134

Visual inspection of the distributions (see Fig 2.2a) indicated three possible changes within Conduct referrals, which were examined with post-hoc comparisons pooling all other categories (corrected critical value of $\alpha=.0169$). These apparent changes were i) a greater number of Black or Black British children were referred in the most recent time period compared with the previous time periods. This was confirmed in the post-hoc comparison ($\chi^2=8.654$, $p=.015$, $df=2$); ii) a reduction in the amount of unavailable or missing data for the two most recent time periods ($\chi^2=12.510$, $p=.002$, $df=2$); and iii) an increase in referrals for children categorised as having mixed ethnicity. However, this final observation was not significant upon post-hoc testing (Fisher's exact=3.081, $p=.233$, $df=2$).

Within A&F cases (see Fig. 2.2b), the difference appeared to stem from i) an increased number of White children referred during Time Period 2 than the other time categories, which was confirmed in a subsequent post-hoc comparison ($\chi^2=15.073$, $p<.001$, $df=2$); ii) a progressive reduction in missing or unavailable information over the time periods studies, also confirmed statistically ($\chi^2=15.897$, $p<.001$, $df=2$); and iii) a possible reduction in the number of mixed race children referred in the most recent time period, which not confirmed ($\chi^2=2.361$, $p=.330$, $df=2$).

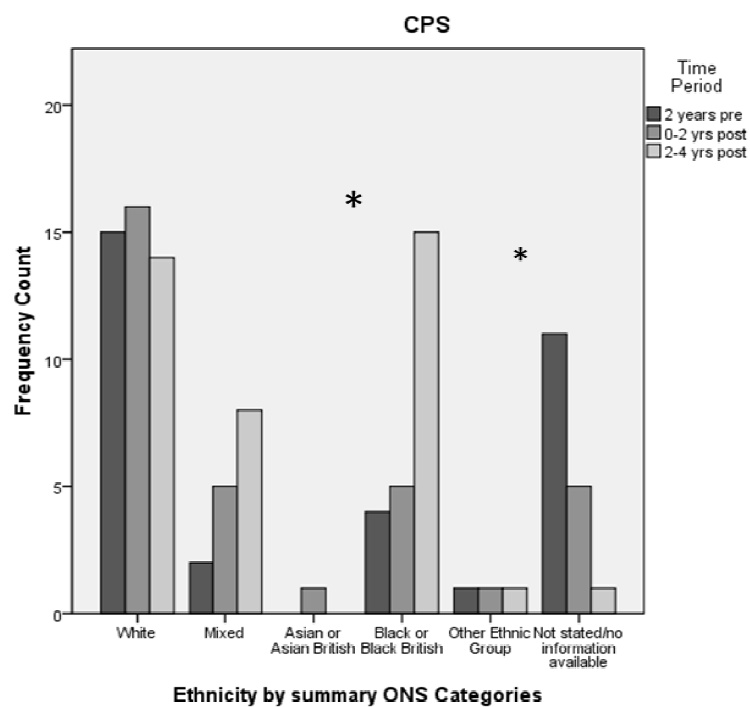
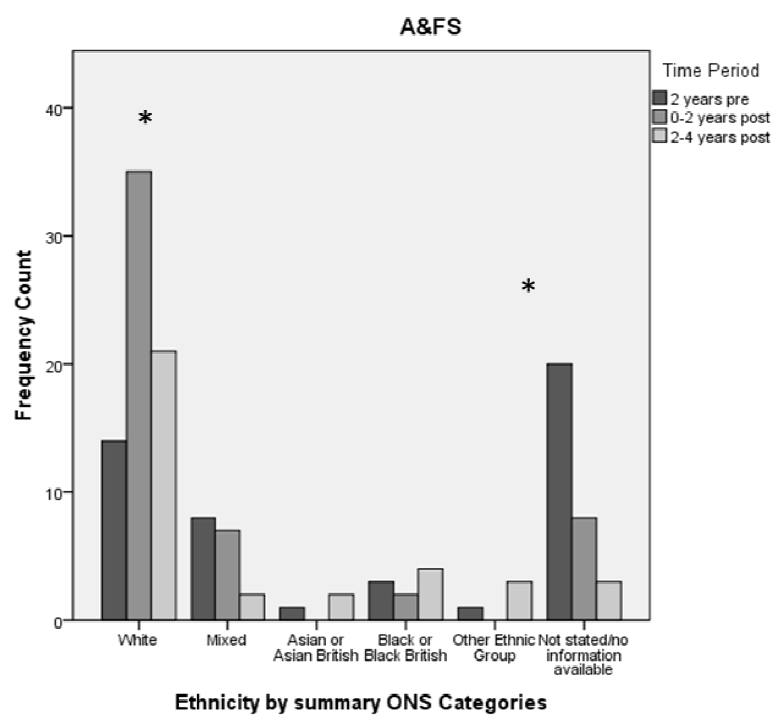
A**B**

Figure 2.2. Distribution of Ethnicity by Time Period within the Conduct Problems Clinic (Panel A) and the Adoption and Fostering Clinic (Panel B).

Asterisks denote changes that were statistically significant ($p < .05$).

4.1.4. Age at Referral

The variable age at referral (in years) was entered as a dependent variable in a univariate ANOVA with the factors of Time Period and Service. There was no main effect of Time Period ($F(2,229)=.578$, $p=.562$), but there was a main effect of Service ($F(1,229)=4.376$, $p=.038$), and a significant interaction between the two factors ($F(2,229)=3.714$, $p=.026$). Inspection of the descriptive statistics in Table 2.4 and of Figure 2.3 shows that this interaction was such that the age of CPS cases reduced over time, whereas the age of A&FS cases remained the same or increased slightly over the same time period. This may indicate that referrals have become more appropriate over time (as the referral criteria for the CPS specifies a 3-8 year age bracket), or a trend towards earlier diagnosis associated with increasing awareness of and availability of effective interventions for conduct problems.

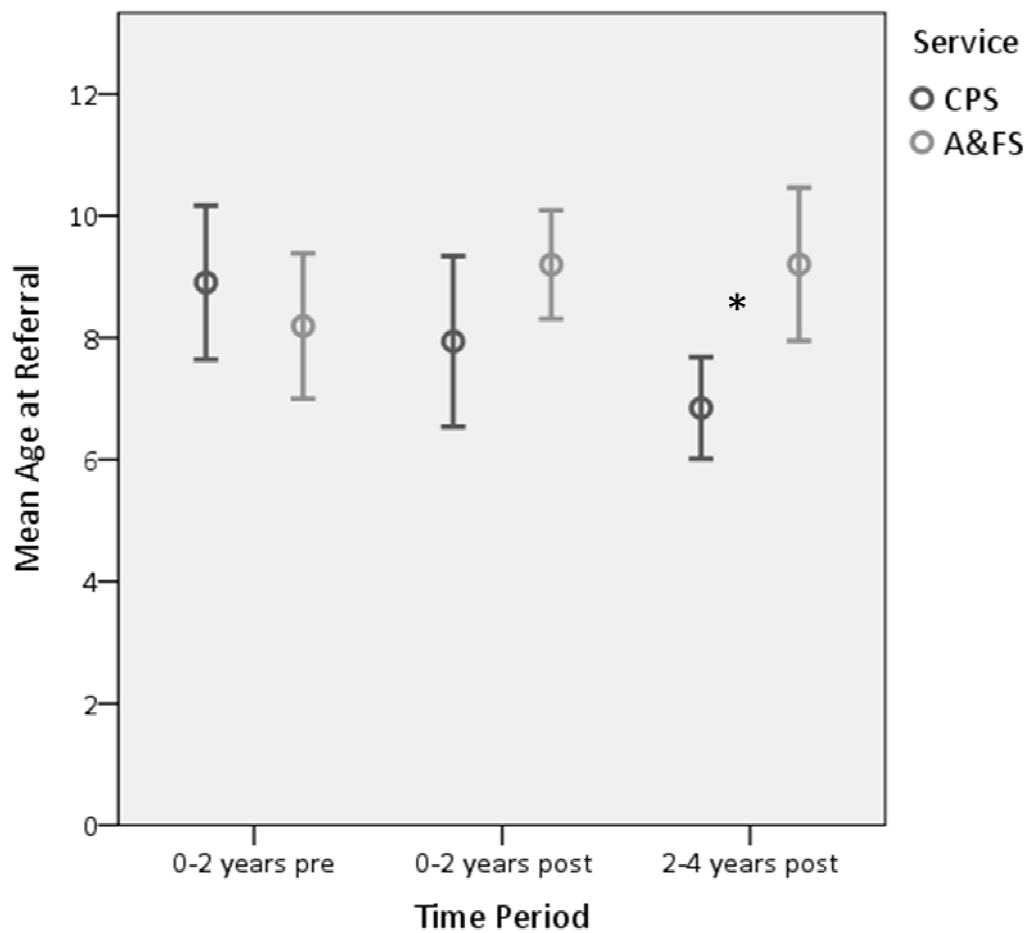


Figure 2.3. Chart showing the age at referrals of cases within each service at each time period.

Error bars are 95% Confidence Intervals (CIs), and the asterisk indicates the source of the interaction, i.e. where there is no overlap between the error bars.

Table 2.4. The mean age (in years) of referrals to each service in each time period, with standard deviations in parentheses.

Service	Time Period		
	2 years pre	0-2 years post	2-4 years post
CPS	8.91 (3.504)	7.94 (3.952)	6.85 (2.570)
A&FS	8.19 (4.057)	9.20 (3.136)	9.21 (3.599)

4.1.5. Parental Mental Health Problems and Learning Disability

There was a significant difference in the distribution of parental mental health problems or learning disabilities over the three time periods in the Conduct Clinic (Fisher's exact=15.053, $p=.016$, $df=6$). Within A&F referrals, the distributions did not differ significantly either for birth parents (Fisher's exact=8.209, $p=.205$, $df=6$) or adoptive parents (Fisher's exact=3.407, $p=.841$, $df=6$).

Table 2.5. The occurrence of any parental mental health problem or learning disability, according to clinic and time period.

Note that missing values are stated for information but were not included as a variable for analysis.

Service	Parental MHP or LD	Time Period			Total
		2 years pre	0-2 years post	2-4 years post	
CPS	None	10	5	4	19
	Mother only	4	8	19	31
	Father only	5	4	2	11
	Both Parents	12	15	12	39
	Total	31	32	37	100
	Missing Data	2	1	2	5
A&FS: Birth Parents	None	1	3	4	8
	Mother only	18	13	13	44
	Father only	1	5	2	8
	Both Parents	19	24	10	53
	Total	39	45	29	113
	Missing Data	7	5	6	18
A&FS: Adoptive Parents	None	20	30	19	69
	Mother only	7	9	5	21
	Father only	2	5	1	8
	Both Parents	1	0	0	1
	Total	30	44	25	99
	Missing Data	16	6	10	32

Inspection of the frequencies in Table 2.5 indicated that the change within the CPS principally reflected an increase in maternal mental health problems or learning disabilities over the time periods studied confirmed statistically ($\chi^2=12.449$, $p=.002$, $df=2$, Sidak-adjusted $\alpha=0.025$), and possibly a reduction in the number of cases with no parental mental health problems or learning disability, though this was not confirmed statistically ($\chi^2=5.390$, $p=.061$, $df=2$, Sidak-adjusted $\alpha=0.025$).

As the category in question is a broad one, finer-grained analyses were conducted on the disorders known to moderate treatment response, depression, and anxiety (NICE, 2006)²⁰. There was a difference in parental depression across the three time periods (Fisher's exact=32.137, $p<.001$, $df=6$), with Figure 2.4a illustrating that this meant more mothers with depression in the most recent time period, and this being confirmed statistically ($\chi^2=19.581$, $p<.001$, $df=2$). Additionally, there was a significant difference in parental anxiety disorders across the time periods (Fisher's exact=11.996, $p=.018$, $df=6$), with Figure 2.4b indicating there were more mothers with anxiety in the two most recent time periods, which was confirmed statistically ($\chi^2=7.265$, $p<.025$, $df=2$).

Note that the 81 cases of mood disorder comprised almost exclusively depression, but included 12 cases of depression and postnatal depression, 1 x postnatal depression only, 1x bipolar, and 1x psychotic depression). There were 20 cases of anxiety disorders, which included 4 cases of panic disorder, 6 of Obsessive Compulsive Disorder, 3 of Post-Traumatic Stress Disorder, and 9 where an anxiety disorder was recorded not otherwise specified. Note two cases had two comorbid anxiety disorders. It is interesting to note that 19 of the 20 cases of parental anxiety were also associated with parental depression.

[See Figure 2.4, overleaf]

²⁰ Other mental health and learning disabilities were recorded, including psychosis, Autism Spectrum Disorders, and rarer mental health problems such as Munchhauens's syndrome, but their relative infrequency precluded any meaningful analysis.

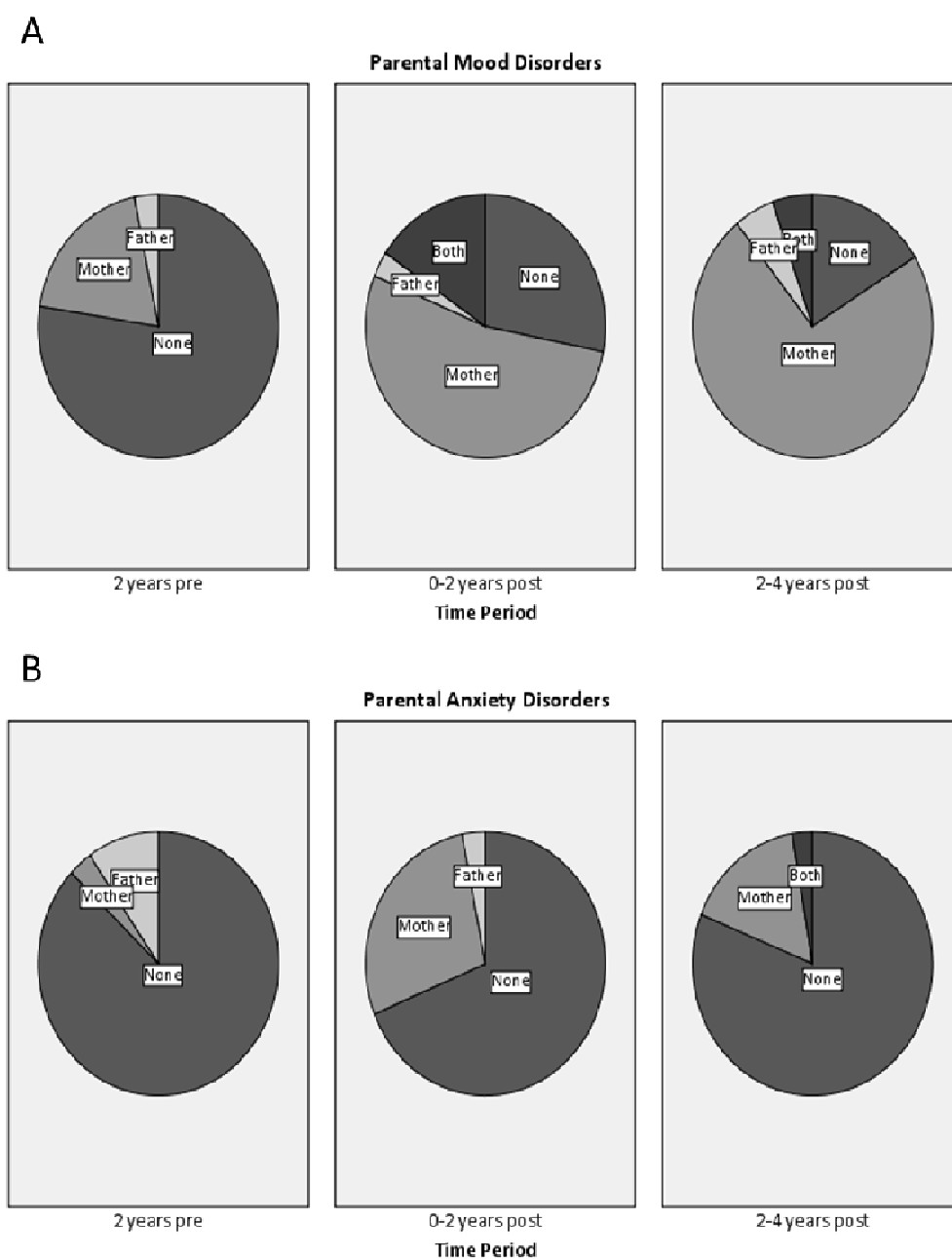


Figure 2.4. Pie charts showing the change in parental mental health problems in the CPS over the three time periods.

Panel A refers to mood disorders, and Panel B Anxiety disorders. In both cases, the occurrence of maternal mental health problems increased over time, though maternal anxiety was most frequent during the middle time period.

4.1.6. *Comparison of Child Factors over the Three Time Periods*

Analyses were conducted to examine any changes in child factors between services over time, including the number and nature of the problems. The number of diagnoses and presenting problems were examined with ANOVAs on the fixed factors of Service and Time Period on the dependent variables of the number of: a) referral diagnoses, b) referrer-identified problems, c) presenting problems, d) outcome diagnoses.

The ANOVA on *number of referral diagnoses* showed a significant effect of service ($F(1,231)=4.884$, $p=.022$, partial eta squared=.023, 63% power), such that there were fewer diagnoses in the A&FS compared with the CPS (.88 vs .58 diagnoses), but no significant effect of time ($F(2,231)=.667$, $p=.485$, partial eta squared=.006, 17% power), and no interaction ($F(2,225)=.020$, $p=.980$, partial eta squared=.000, 5% power).

The ANOVA on the *number of referrer-identified problems* showed significant effects of service ($F(1,225)=8.910$, $p=.044$, partial eta squared=.018, 52% power) with more problems identified in the A&FS than the CPS (2.73 problems versus 2.25); a significant effect of time ($F(2,225)=13.537$, $p=.002$, partial eta squared=.052, 89% power), with fewer problems in the most recent time period (2.58, 2.90, and 2.0), but no interaction ($F(1,231)=4.884$, $p=.022$, partial eta squared=.023, 63% power). This pattern was mirrored in the ANOVA on team-defined presenting problems and so the results will not be repeated.

The ANOVA on the *number of post-assessment diagnoses* showed a significant effect of Time ($F(2,222)=3.920$, $p=.021$, partial eta squared=.034, 70% power), with fewer diagnoses in the more recent time period (1.14, 1.31, and .91 diagnoses for the three time points respectively), a trend in Service ($F(1,222)=3.542$, $p=.061$, partial eta squared=.016, 47% power) with more diagnoses in the CPS than A&FS (1.24 versus 1.03), with no interaction ($F(2,222)=1.851$, $p=.159$, partial eta squared=.016, 38% power).

No statistically significant differences were observed in the types of referral diagnoses, referrer-identified problems, team-identified problems, parent- or teacher-rated SDQ scores, or the majority of the psychometric variables (see Appendix 2). There were, however, significant differences in the number of children in the A&FS diagnosed with CD/ODD (Fisher's exact = 11.765, $p=.029$, $df=4$), ADHD ($\chi^2=12.207$, $p=.015$, $df=4$), Intellectual Disability (Fisher's exact = 10.266, $p=.017$, $df=4$), and Specific Disorder or Spelling (Fisher's exact=14.348, $p=.001$, $df=4$) over the three time periods. In each case this stemmed from there being more diagnoses made in the middle time period (see Table 2.6). As would be expected, some psychometric assessment indices were accordingly worse in this group and time period (see Appendix 2).

Table 2.6. Diagnostic frequencies within the A&FS over the three time period.

Only those diagnoses that differed significantly over time are shown.

	2 years pre	0-2 years post	2-4 years post
CD/ODD	21	31	14
ADHD	8	20	4
Intellectual Disability	5	10	1
Specific disorder of spelling	3	11	2

Since the guideline was published, the CPS has seen an increase in CAMHS referrals, referrals regarding Black/Black British families, and the number of cases with associated maternal mental health problems, along with decreases in age of referrals, and a decrease in the number of presenting problems and diagnoses.

The changes in age and ethnicity of referrals do not seem to be linked with complexity or treatment failure in any meaningful way (see discussion). The increase in maternal mental health problems and in past contact with mental health services are both factors associated with case complexity, but the reduction in the number of presenting problems and diagnoses conversely indicates a potential reduction in case complexity. The case complexity index therefore incorporated all of these factors with the aim of illuminating this issue.

4.2. Examination of Complexity Index by Time Period

The complexity index was entered as the dependent variable into a univariate ANOVA on the factors of Time Period (3 levels: pre, post1, post2), and Service (two levels, CPS and A&FS). The main effect of Time Period was significant ($F(2,221)=3.554$, $p=.030$, partial eta squared = .031, 66% power). This indicated that the complexity of cases seen changed over the time period studied, with greater complexity in the later time points. The main effect of Service was also significant ($F(1,221)=16.886$, $p<.001$, partial eta squared = .071, 98% power), such that there were more complex cases in the CPS compared with the A&FS. The interaction between Time Period and Service was not statistically significant ($F(2,221)=1.774$, $p=.172$, partial eta squared=.016, 40% power), indicating that the increase in complexity over time did not differ between the two clinics. However, this analysis was under-powered.

Visual inspection of the data in Figure 2.5, paying particular attention to the 95% Confidence Intervals, does indicate that the patterns for each service differ over time. Specifically there is almost complete overlap in the error bars for the CPS and A&FS in the first time point, and yet there is no overlap at all between the services in the subsequent two time periods.

Indeed, when the time categories were collapsed into simple 'pre-guideline' and 'post-guideline' categories (to increase statistical power), the interaction term was statistically

significant ($F(1,123)=4.298$, $p=.049$, partial eta squared=.017, 51% power). This interaction was such that A&FS complexity remained stable over time, whereas the CPS complexity increased over the same period. This demonstrates that despite the variation in the change seen in constituent complexity factors over the time periods, overall complexity has increased in the cases seen in the CPS following the publication of the TA, and the same change is not evident in the comparison cases from the A&FS. See also Table 2.19 for descriptive statistics.

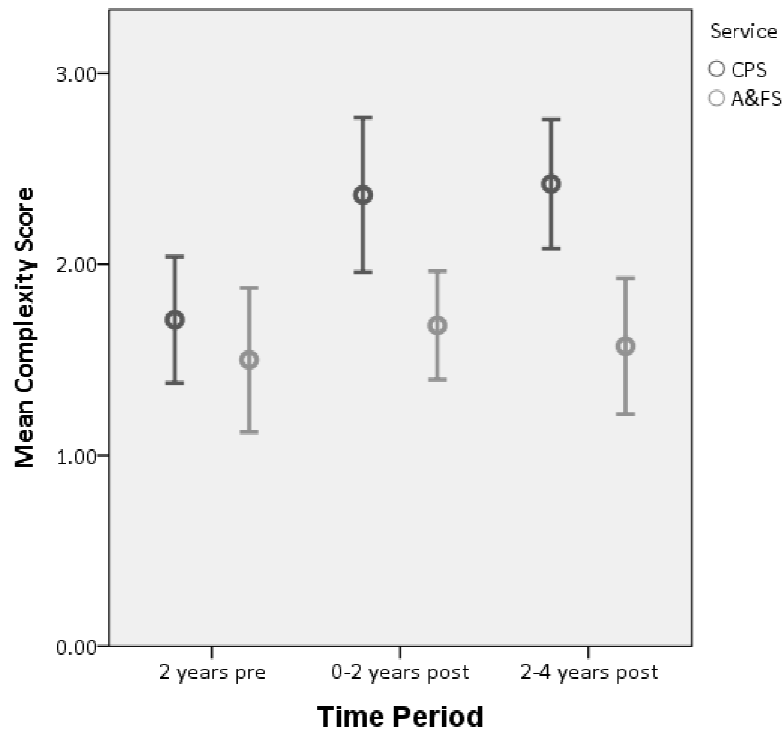


Figure 2.5. Composite Complexity Index Scores according to Service and Time Period.

Error bars show 95% CIs.

4.3. Examination of Treatment Plans by Time Period

4.3.1. Number of Treatments

The next phase of the analysis was to examine the impact of the increase in complexity on the service, in terms of the number and range of treatments offered. This involved a univariate ANOVA on the number of treatments for each case, according to the factors of service and time period. There was a trend towards a main effect of time ($F(2,226)=4.074$, $p=.088$, partial eta squared=.021, 49% power) in the direction that more treatments were planned after the guideline, but no effect of service ($F(1,226)=.084$, $p=.772$, partial eta squared, .00 6% power), and no interaction ($F(2,226)=1.160$, $p=.315$, partial eta squared=.010, 10% power). Although

the analysis was under-powered, it does mirror the pattern seen in the complexity index for the CPS seen in Figure 2.5, albeit to a reduced degree.

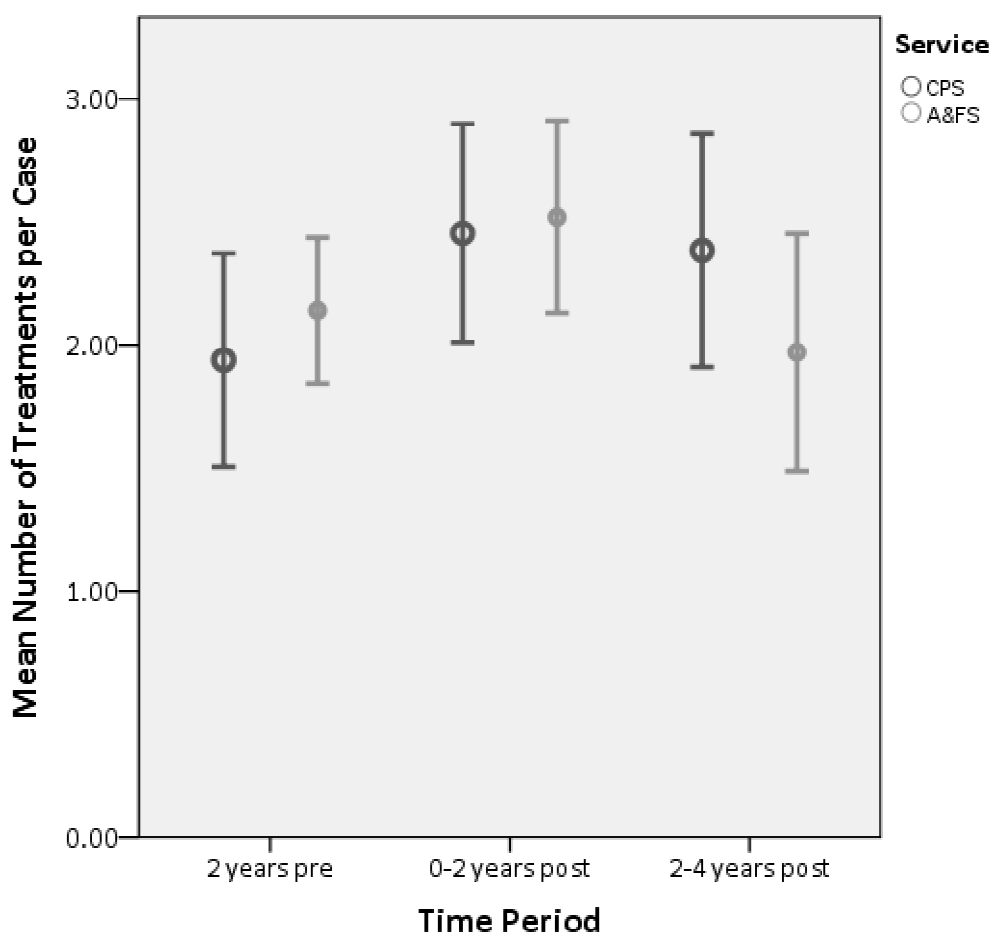


Figure 2.6. The number of treatments planned per family by service and time period.

Error bars show 95% CIs.

4.3.2. Examination of Treatment Plans by Time Period

The next stage of the analysis involved the examination of contingency tables of the most common components of treatment plans, which paralleled the previously reported analyses of categorical variables. The components in question were PCG, CBT, anger management, social skills training, life story work, social services liaison, school liaison and further psychometric assessment²¹. Only those components showing significant change are reported here.

²¹ Note that there was no overall change in the frequency with which further psychometrics were carried out over the time. However, the number of ASD assessments conducted in-house increased from zero in the first time period, to one in the middle time period, to six in the most recent time period. There was a concomitant decrease in the number of referrals for external assessment ($n=5$ at time 1,

4.3.2.1. Parent-Child Game

There was a significant difference in the proportion of cases offered Parent-Child Game as part of the treatment plan over the time periods, within the CPS ($\chi^2=22.527$, $p<.001$, $df=2$). Inspection of Table 2.7 showed that this involved a clear increase in cases offered the treatment over time.

Table 2.7. Percentage of cases with PCG as part of the treatment plan over the three time periods and two clinics

Service	Time Period		
	2 years pre	0-2 years post	2-4 years post
CPS	24.2%	45.5%	79.5%
A&FS	34.0%	32.0%	23.5%

4.3.2.2. Anger Management

There was also a significant difference in the number of cases where anger management was a planned treatment, in the CPS (Fisher's exact=6.480, $p<.025$, $df=2$). Inspection of Table 2.8 showed that more cases were offered anger management in the middle time period.

Table 2.8. Percentage of cases with Anger Management as part of the treatment plan by time period and clinic.

Service	Time Period		
	2 years pre	0-2 years post	2-4 years post
CPS	0%	15.2%	2.6%
A&FS	6.8%	2%	8.6%

4.3.2.3. Behavioural Family Work

There was also a significant difference in the number of cases where behavioural family work was a planned treatment, within the A&FS (Fisher's exact=6.149, $p<.046$, $df=2$). Inspection of Table 2.9 showed that more cases were offered this work in the middle time period.

Table 2.9. Percentage of cases with Behavioural Family Work as part of the treatment plan by time period and clinic.

Service	Time Period		
	2 years pre	0-2 years post	2-4 years post
CPS	18.2%	18.2%	15.4%
A&FS	13.6%	36.0%	25.7%

4.3.2.4. Medication

There was also a significant difference in the number of cases where medication was a planned treatment, in the A&FS ($\chi^2=13.933$, $p<.001$, $df=2$). Inspection of Table 2.10 showed that more cases were offered medication in the middle time periods. The vast majority of prescriptions were for medications for ADHD (e.g. methylphenidate, atomoxetine) and sleep problems (melatonin).

n=6 at time 2, n=2 at time 3), reflecting increased capacity and competency to complete such assessments within the team. However, these changes did not reach statistical significance.

Table 2.10. Percentage of cases with Medication as part of the treatment plan by time period and clinic.

Service	Time Period		
	2 years pre	0-2 years post	2-4 years post
CPS	30.3%	33.3%	23.1%
A&FS	11.4%	46.0%	25.7%

4.3.2.5. School Liaison

There was also a significant difference in the number of cases where school liaison was a planned treatment, in both the CPS ($\chi^2=7.934$, $p<.018$, $df=2$) and the A&FS ($\chi^2=8.439$, $p<.015$, $df=2$). Inspection of Table 2.11 showed that within the CPS there were fewer cases whose treatment plans included school liaison than the previous two time periods. Within the A&FS, there were more cases involving school liaison in the middle time point.

Table 2.11. Percentage of cases with Medication as part of the treatment plan over the three time periods and two clinics.

Service	Time Period		
	2 years pre	0-2 years post	2-4 years post
CPS	40.4%	51.5%	20.5%
A&FS	31.8%	56.0%	28.6%

5. Discussion

5.1. Summary of Results and their Relation to the Hypotheses

The main finding of this audit is that the complexity of referrals accepted by a Tier 4 Conduct Problems Service increased in the four-year period following the publication of the NICE Technology Appraisal, relative to the 2 previous years. There was no such change in complexity of referrals accepted by a linked Adoption and Fostering Service over the same time period. Complexity was indexed by a composite measure incorporating presence of maternal mental health problems, previous contact with mental health services, comorbid disorders, borderline intellectual functioning, and impaired literacy. A series of finer-grained analyses were broadly consistent with the main result. Specifically, the findings were that:

- In the most recent time period, there was a large increase in the proportion of referrals to CPS from CAMHS relative to other sources (predominantly general practitioners). This indicates that more cases have had previous contact with mental health services and consequently are more likely to have had a poor response to, or difficulty engaging with parenting programs. Specific data on treatment history was, however, not available.
- Within the CPS, there were more referrals accepted regarding children of Black or Black British ethnicity in the most recent time period, and in the A&FS there were more referrals accepted for children classified as having a White ethnic background in the middle time period. In both CPS and A&FS groups, there were fewer missing data on ethnicity in more recent time periods, which may reflect improvements in clinical data management procedures.
- The average age at referral reduced over time in the CPS, but not in the A&FS. This may reflect a benefit of the TA, that increased awareness of effective interventions and improved management of these cases at Tiers 2-3 has prompted earlier referrals for treatment. There may also be an influence of improved marketing for the A&FS, making clear that the service works with young people up to the age of 18 years.
- Within the CPS, there were more cases where mothers had mental health problems or learning disabilities in the most recent time period. This change reflected an increase in both mood and anxiety disorders. There were no such changes evident in the A&FS. As maternal depression, anxiety and general self-reported stress is important in determining treatment outcomes (Kazdin, 1990; NICE, 2006), this is an important finding.

- There has been no meaningful change in either the proportion of cases referred with, or subsequently given a diagnosis over the groups and time periods studied, or in the type of diagnoses given (aside from some change specific to the mid-period in the A&FS which does not seem relevant for the purposes of this audit). However, it does appear that there was a slight reduction in the number of presenting problems and consequent diagnoses in both services in the most recent time period.
- The team had hypothesised that there may be an increase in the number of comorbid diagnoses, specifically ADHD and ASD, across both the CPS and A&FS, but this hypothesis was not supported. There were however, numerical trends towards increased in-house assessment of ASD and reduced external referrals for such assessments in the most recent time period.
- There were no apparent changes in the psychometric characteristics of cases accepted by either the CPS or A&FS over the three time periods. Nor were there any changes in parent- or teacher-ratings on the Strengths and Difficulties Questionnaire. However it should be noted that statistical power for some of these analyses was low.

Taken together these detailed analyses of the complexity data indicate relative stability in terms of the children's presenting problems, but an increase in complexity stemming from maternal mental health problems and possible treatment non-completion or non-response.

The impact of this increase in case complexity was hypothesised to have had an impact on the complexity of treatment plans. This hypothesis was not clearly supported, as only a trend towards increased numbers of interventions was observed, and this was across both services. However, in the CPS, there was a large increase in the number of families offered the Parent-Child Game intervention after the TA, and more children were offered Anger Management training within the middle time period but this was not maintained.

Within the A&FS, more children were offered behavioural family interventions, medication, and school liaison during the middle time point, compared with the other time points. This seems likely to be related to the increased variety of diagnoses present in cases from this time period, however any further emergent trends will need to be explored in future audits.

The overall picture in terms of the impact of increased complexity on treatment plans seems to be that there is a slightly smaller variety of treatments offered, but a large increase in the proportion of cases offered the specialist PCG intervention.

5.2. Complexity and the TA: Implications for Families and Services

The increases in CPS case complexity and in the number of PCG interventions offered since July 2006 imply that the recommendations from the TA are being followed. This has implications for families accessing services as well as for CAFT and broader CAMHS and health service commissioning.

For families, this should mean that effective interventions can be accessed more easily and quickly than has previously been the case. The current expansion of the Improving Access to Psychological Therapies initiative to children and young people would be expected to further increase the availability. In the longer term, such initiatives may also serve to reduce the stigma still associated with common mental health problems. However, for complex cases, such programmes may be insufficient, and this may mean that access to the necessary specialist approaches to parenting for those with the greatest need may be delayed, and there may be increased probability of experiencing treatment failures. It is therefore important that cases likely to need interventions more complex than standard group-based programmes are identified and referred to teams such as CAFT quickly and efficiently. This would require competence within local CAMHS to make decisions on when to 'step up' referrals to Tier 4. The reducing age of CPS referrals over time, however, fits with the more optimistic hypothesis.

There are implications for the CAFT team related to the effective identification of appropriate referrals. For example, it can be difficult to establish on the basis of a brief letter from a GP whether a case meets criteria for individual PCG. Perhaps a referral form including checkboxes to indicate potential sources of complexity may facilitate such identification. The increasing complexity of referrals also has more immediate clinical implications. Working with families who are difficult to engage, for whom previous treatments may have been unsuccessful, who may be experiencing a variety of socio-economic stresses, and include parents with their own mental health difficulties, presents a real challenge. The high incidence of maternal mental health problems in particular indicates that there may be a role for an adult mental health worker within the team, to complement and facilitate parenting interventions. Innovation in clinical practice is necessary to ensure the best outcomes. However, even relatively simple adaptations such as increasing the number of sessions (as CAFT have done, increasing from 6-8 to 8-10 sessions), having flexible clinic hours, and offering home visits can have large resource implications. Given the ongoing cuts to NHS funding this is likely to become even more of a challenge with time. Finally, the increasing availability of PCG programmes in non-specialist services may also necessitate an increased role for specialists within CAFT to provide training and consultation to colleagues in other service settings.

5.3. The Impact of the TA on the A&FS

The impact of the TA on the A&FS was much less clear than its impact on the CPS. This clearly fits with the team's expectations in light of the previously discussed conceptualisation of conduct problems in children who are adopted or fostered within attachment frameworks. For the purposes of this audit, data from the A&FS has served as a useful form of quasi-control group. However, it is worrying that there has been so little impact of the guidance on clinical practice in this area, as it implies that there has been no change in the number of adopted or fostered children participating in PCG programmes, which we know to be extremely effective. As forthcoming NICE guidance documents will specifically refer to interventions for this population, the situation should improve.

5.4. Suggestions for Further Investigation

There are many potential avenues in which further audit and research may progress. One of the foremost questions raised by the current findings is whether or not the increase in complexity has brought about any change in clinical outcomes, and more broadly whether there is an impact of complexity on treatment outcomes, both within the CPS and the A&FS. In due course, it will also be interesting to monitor changes in A&FS referrals following the publication of NICE guidance that includes this group. Additionally, it would be interesting to explore service user perspectives on satisfaction with the service, improving accessibility, or removing barriers towards engagement.

A more detailed investigation of treatment plans could also be useful, as the present analysis only concerned treatment plans made immediately after initial assessment, and the number of treatments planned was the only index of complexity. As it stands in the CPS, the reduction in treatment variety and increase in the number of cases offered PCG is consistent with the observed decrease in the number of presenting problems and diagnoses. Furthermore there was an increase in the variety of treatments planned in the A&FS in the middle time-period, when there was also an increase in the variety of diagnoses. These findings suggest that the variety of the treatment plan is closely linked to the variety of the child's presentation. However, there may be other influences on treatment plans not considered here, such as important advances in the evidence base for some treatments, or alterations in staff capacity to offer particular interventions. Furthermore, these analyses do not take into account modifications to the treatment plan, treatments completed versus those planned, or importantly, adaptations to PCG. Such adaptations include having more sessions, including cognitive work with parents, collaborations with family support workers etc, which are

necessary given the greater numbers of families where parents have mental health problems and/or who have previously been involved with CAMHS.

Future projects would be facilitated if data on important complexity variables other than those recorded here were recorded in a systematic fashion. For example, referral information was often difficult to find, and referral letters rarely stated whether or not the child had any existing diagnoses, and if so what they were. It is possible that the use of a standard referral form with spaces to indicate existing diagnoses, the referral problem, and checklists regarding important matters related to complexity such as parental mental health issues and history of parent training interventions, might help to efficiently direct referrals to the most appropriate team as well as facilitating future audits. However, it is acknowledged that referrers may not welcome additional paperwork, and it is important to make the referral process as user-friendly as possible.

5.5. Limitations of the Current Project

This audit has several methodological limitations. The focus of this report has been on changes in referrals following publication of the TA, however, it cannot be stated that the publication of the TA actually caused the changes observed. Although the A&F referrals act as a quasi-control group here, there may be other organisation and economic factors that could have influenced referrals to the CPS but not the AFS. Also in relation to the use of a quasi-control group, this was obviously not a randomised procedure, the groups were unbalanced in terms of age and gender, and there may have been other systematic differences between the groups. Furthermore, the separation of the data into two year groups is somewhat arbitrary; there is unlikely to be any difference between a referral from July 2008 and August 2008, but they are represented in different categories. Nonetheless, this approach is a logically sound means of capturing change over time.

There were some variables that it would have been interesting to compare, but for which data were not available from the existing dataset, specifically the presence of child protection concerns at referral, or the necessity of making a child protection referral, as an important index of complexity. Additionally, only the post-assessment treatment plans were analysed, and it would have been interesting to include all treatments carried out up to the point of discharge. In this way, the analyses here may not have captured the complexity of presentations which may have emerged and subsequent interventions offered by the team.

The reliability and validity of the case-note review procedure is not known. For some variables, there data are incontrovertible (e.g. age, gender), but in others there is more room for interpretation (e.g. presenting problem). There is the possibility that relevant information was

missed or inaccurately recorded during initial recording and in the process of retrieving information from files. Although this sort of variation due to human error is expected in any dataset, and should be randomly distributed, there is a source of potential bias in the current project as a different researcher retrieved the data for the third time point than the first and second.

Finally, there are some limitations to the statistical analysis carried out. Specifically, the observed power for the ANOVA on parent and teacher-rated SDQs was low, and as such there is a risk of a Type 1 error (i.e. incorrectly accepting the null hypothesis). Much more data would be needed to convincingly demonstrate the null result of there being no change in subjectively-rated problems over the time period studied. There is also a limitation related to the use of chi-square and Fisher's exact tests. Though these are appropriate tests to use for this analysis, they do not allow for examination of the interaction term as in methods such as ANOVA. Specifically, where there is a significant effect in the CPS data and no significant difference in the A&FS data, it does not necessarily mean that the difference between CPS and A&F is itself a statistically significant one (i.e. the non-significant effect may just might be underpowered). This is sometimes described in terms of the difference between $p=.049$ and $p=.051$ not being significant (e.g. Nieuwenhuis, Forstmann, & Wagenmakers, 2011). Nonetheless, the analysis remains appropriate for the data and current purpose.

6. References

- Ainsworth, M. D. (1979). Infant-mother attachment. *American Psychologist*, 34(10), 932–937.
- Baldock, E. (2008). *Audit of the NICE Technology Appraisal for conduct disorder, and study of the prevalence and treatment of mental disorders in adopted and fostered children*. (Doctoral Dissertation in Clinical Psychology). King's College London.
- Bandura, A. (1976). *Social Learning Theory* (1st ed.). London: Pearson.
- Bohman, M. (1996). Predisposition to criminality: Swedish adoption studies in retrospect. *Ciba Foundation Symposium*, 194, 99–109; discussion 109–114.
- Booth, P. B., & Jernberg, A. M. (2010). *Theraplay: Helping Parents and Children Build Better Relationships through Attachment-Based Play* (3rd Edition.). John Wiley & Sons.
- Bowlby, J. (1953). *Child Care and the Growth of Love*. Middlesex, England: Pelican Original.
- Bywater, T., Hutchings, J., Linck, P., Whitaker, C., Daley, D., Yeo, S. T., & Edwards, R. T. (2010). Incredible Years parent training support for foster carers in Wales: a multi-centre feasibility study. *Child Care, Health and Development*, 37(2), 233–243. doi:10.1111/j.1365-2214.2010.01155.x
- Carr, A. (2006). *The Handbook of Child and Adolescent Clinical Psychology: A Contextual Approach* (2nd ed.). Hove, East Sussex: Routledge.
- Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., ... Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297(5582), 851–854. doi:10.1126/science.1072290
- Dretzke, J., Davenport, C., Frew, E., Barlow, J., Stewart-Brown, S., Bayliss, S., ... Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: a systematic review of randomised controlled trials. *Child and Adolescent Psychiatry and Mental Health*, 3(1), 7. doi:10.1186/1753-2000-3-7
- Fisher, P. A., & Kim, H. K. (2007). Intervention effects on foster preschoolers' attachment-related behaviors from a randomized trial. *Prevention Science: The Official Journal of the Society for Prevention Research*, 8(2), 161–170. doi:10.1007/s11121-007-0066-5
- Friedli, L., & Parsonage, M. (2007). *Mental Health Promotion: Building an Economic Case*. Belfast: Northern Ireland Association for Mental Health.
- Hughes, D. (2004). An attachment-based treatment of maltreated children and young people. *Attachment & Human Development*, 6(3), 263–278. doi:10.1080/14616730412331281539
- Kazdin, A. E. (1990). Premature Termination from Treatment among Children Referred for Antisocial Behavior. *Journal of Child Psychology and Psychiatry*, 31(3), 415–425. doi:10.1111/j.1469-7610.1990.tb01578.x

- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney: Psychology Foundation.
- McMahon, R. J., & Forehand, R. L. (2005). *Helping the Noncompliant Child, Second Edition: Family-Based Treatment for Oppositional Behavior* (2nd ed.). New York: Guilford Press.
- Minnis, H., Pelosi, A. J., Knapp, M., & Dunn, J. (2001). Mental health and foster carer training. *Archives of Disease in Childhood*, 84(4), 302–306.
- Moffitt, T. E. (2005). Genetic and environmental influences on antisocial behaviors: evidence from behavioral-genetic research. *Advances in Genetics*, 55, 41–104. doi:10.1016/S0065-2660(05)55003-X
- NICE. (2006). *Conduct Disorder in Children - Parent-Training/Education Programmes* (Guidance/Technology Appraisals). London: National Institute of Health and Clinical Excellence. Retrieved from <http://www.nice.org.uk/>
- NICE. (2010). *Conduct Disorders in children and young people stakeholder consultation workshop* (Consultation Report). London: NICE. Retrieved from <http://guidance.nice.org.uk/CG/Wave24/8/Scoping>
- Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature Neuroscience*, 14(9), 1105–1107. doi:10.1038/nn.2886
- Nilsen, W. J. (2003). Perceptions of attachment in academia and the child welfare system: the gap between research and reality. *Attachment & Human Development*, 5(3), 303–306. doi:10.1080/14616730310001593910
- Pallett, C., Scott, S., Blackeby, K., Yule, W., & Weissman, R. (2002). Fostering changes: a cognitive-behavioural approach to help foster carers manage children. *Adoption & Fostering Journal*, 26(1), 39–48.
- Rushton, A., Monck, E., Leese, M., McCrone, P., & Sharac, J. (2010). Enhancing adoptive parenting: a randomized controlled trial. *Clinical Child Psychology and Psychiatry*, 15(4), 529–542. doi:10.1177/1359104510365041
- Scott, S. (1998). Aggressive behaviour in childhood. *British Medical Journal (Clinical Research Ed.)*, 316(7126), 202–206.
- Simmel, C., Brooks, D., Barth, R. P., & Hinshaw, S. P. (2001). Externalizing symptomatology among adoptive youth: prevalence and preadoption risk factors. *Journal of Abnormal Child Psychology*, 29(1), 57–69.
- SPSS Inc. (2009). *PASW Statistics 18*. Chicago, IL: IBM.
- Tapsfield, R., & Collier, F. (2005). *The Cost of Foster Care: Investing in our Children's Future*. London: British Association for Adoption and Fostering.

Yates, P., Garralda, M. E., & Higginson, I. (1999). Paddington Complexity Scale and Health of the Nation Outcome Scales for Children and Adolescents. *The British Journal of Psychiatry*, 174(5), 417–423. doi:10.1192/bjp.174.5.417

7. Appendix 1: Data Collection Form

CAFT Database Record Form

Date referred: _____ ID Code: _____

Date of Birth: _____ Age: _____ Sex: M F Ethnicity: _____

Referral Source: GP / T2 / T3 / Other (specify) Local / National _____

Referral Type: Conduct Disorder / A&F / Other (specify) _____

Referral Dx:

None / CD / ODD / ADHD / ASD / Attachment Disorder / Other (specify) _____

Referral Problem:

Conduct Sx / ODD Sx / ADHD Sx / ASD Sx / Mood disorder / Anxiety or PTSD / Enuresis or Encopresis / Risk of placement breakdown / Anger / Attachment difficulties / LD / Other

Source of Info: Referral Letter / Social Services Report / Paediatric Report / Tx Hx / Other

Notes/specify

Presenting Problem (from Part 1):

Conduct Sx / ODD Sx / ADHD Sx / ASD Sx / Mood disorder / Anxiety or PTSD / Enuresis or Encopresis / Risk of placement breakdown / Anger / Stealing / Lying / Attachment difficulties / Neurodevelopmental / LD / Other

Notes/specify

Social Services involvement? Past / Present / No / Not known

PRU: Past / Present / No Specify: _____ SEN: Y / N Specify: _____

Parental Mental Health (Birth)

Maternal: NR / Drugs or Alcohol / Depression / Anxiety / Psychosis / ASD / None / Other

Paternal: NR / Drugs or Alcohol / Depression / Anxiety / Psychosis / ASD / None / Other

Parental Mental Health (Adoptive/Foster)

Maternal: NR / Drugs or Alcohol / Depression / Anxiety / Psychosis / ASD / None / Other

Paternal: NR / Drugs or Alcohol / Depression / Anxiety / Psychosis / ASD / None / Other

Parental Occupation – Healthcare and allied professions: Yes / No (specify)

Notes/specify

Psychometrics

WASI / WPPSI / WISC-IV / WISC-III	Literacy: WIAT-II / WORD	
VIQ / VCI =	Word Reading =	Spelling =
PIQ / PRI =	AA Discrep? (sig .05) Y / N / NR	AA Discrep? (sig .05) Y / N / NR
FSIQ =	Freq:	Freq:
Other Tests		

Questionnaires

Intake SDQ	Teacher	Self	Parent	Diagnostic Pred Risk Low, Medium, High
Overall Stress				Any Dx
Emotional Distress				
Behavioural Diffs				Emotional Dx
HA & Attention				
Peer Relationships				Behavioural disorder
Prosocial				
Impact				HA/concentration disorder
Follow-up recorded? Y / N / n/a				

Placement History (A&F only)

Age taken into care: _____ No. previous foster placements: _____

Start date of current placement: _____ Type of current placement: _____

Assessment Outcome

- ☐ CD firm / possible ☐ ODD firm / possible ☐ SLD: Word Reading
☐ ADHD firm / possible ☐ ASD firm / possible ☐ SLD: Spelling
☐ C-U firm / possible ☐ Other (specify) ☐ Learning Disability
☐ No Dx

For ADHD & ASD Dx only: Raised in Part 1? Yes / No Investigated: Yes / No Confirmed?: Yes / No

CP Referral: At Assessment? Yes / No At Follow-Up?: Yes / No

Post-Assessment Tx Plan**A) In-house**

PCG / CBT / Anger Management / Social Skills Training / Life Story Work / Child-Carer Behavioural Family Work / Inter-adult Behavioural Family Work / Medication (specify) / Further Psychometrics (specify) / Other (specify)

Notes/specify

B) Referral to External Agency: Yes / No

To: Social Services / GP / T2 / T3 / Other (specify) _____

For: Specialist Ax (specify) / Parenting group / CBT / Anger Management / Social Skills Training / Life Story Work / Family Therapy / Medication / Other (specify)

Notes/specify

Discharged to: GP / T2 / T3 / Other (specify) _____

8. Appendix 2: Supplementary Statistical Analyses

To examine whether any change had occurred in the type and number of problems over time, a series of ANOVA and chi square analyses focussed on presence/absence of, and type of diagnoses (on referral and after assessment), presenting problems (defined by referral and after assessment), SDQ scores (parent and teacher ratings), intellectual ability (Verbal, Performance and Full-Scale Indices or equivalents) and literacy (WIAT-II standard scores).

8.1. Presence vs Absence of Diagnosis at Referral and After Assessment

There were no significant differences in the proportion of cases referred with or without an existing diagnosis within either clinic (CPS: $\chi^2=3.439$, $p=.206$, $df=2$; A&FS $\chi^2=4.042$, $p=.146$, $df=2$), or in the proportion of cases given a diagnosis after assessment (CPS: Fisher's exact=5.951, $p=.053$, $df=2$; A&FS Fisher's exact=4.488, $p=.113$, $df=2$). See Table 2.12 for details.

Table 2.12. Percentage of cases with a diagnosis at referral, and a diagnosis after assessment, by time period and clinic.

Service	Diagnoses at Referral			Diagnoses after Assessment		
	2 yrs pre	0-2 yrs post	2-4 yrs post	2 yrs pre	0-2 yrs post	2-4 yrs post
CPS	52%	58%	36%	82%	94%	97%
A&FS	32%	33%	51%	70%	87%	71%

8.2. Diagnostic Frequencies at Referral and Outcome

The most frequently occurring diagnoses were conduct disorder and ADHD within the CPS, and Attachment Disorder and ADHD within the A&FS. As statistical analyses on very rare diagnoses would not be meaningful, such analyses were only conducted for categories with at least 15 relevant cases. However, none of the analyses suggested any significant change over the three time periods studied (see Table 2.13), suggesting that there has been no clear change in the referral diagnoses of cases accepted by either service within CAFT over time.

Table 2.13. A summary of analyses conducted on referral diagnoses by time period and clinic.

Referral Diagnosis	CPS	A&FS
Attachment Disorder	N=0, not analysed	N=18, $\chi^2=2.640$, $p=.296$, $df=2$
CD/ODD	N=34, Fisher's exact =.997, $p=.928$, $df=2$	N=10, not analysed
ADHD	N=23, $\chi^2=3.395$, $p=.201$, $df=2$	N=16, $\chi^2=.039$, $p=1.0$, $df=2$
ASD	N=5, not analysed	N=6, not analysed
Anxiety disorder	N=6, not analysed	N=5, not analysed
Mood disorder	N=5, not analysed	N=3, not analysed
Intellectual Disability	N=0, not analysed	N=3, not analysed
Specific Learning Disability	N=4, not analysed	N=2, not analysed
Enuresis or Encopresis	N=0, not analysed	N=1, not analysed

The equivalent analysis of outcome diagnoses revealed some significant changes, outlined in the main body of the report. The remaining analyses are detailed in Table 2.14.

Table 2.14. A summary of the analyses conducted on the outcome diagnoses for each clinic over time.

Asterisks denote statistically significant results.

Outcome Diagnosis	CPS	A&FS
Attachment Disorder	N=1 diagnosis, not analysed	N=10 diagnoses, not analysed
CD or ODD	N=71, Fisher's exact = 6.880, p=.079, df= 4	N=66, Fisher's exact = 11.765, p=.029, df= 4*
ADHD	N=52, $\chi^2=4.392$, p=.350, df=4	N=63, $\chi^2=12.207$, p=.015, df=4*
ASD	N=17, Fisher's exact = 3.392, p=.561, df= 4	N=11, not analysed
Anxiety Disorder	N=10, not analysed	N=9,, not analysed
Mood Disorder	N=6, not analysed	N=5,, not analysed
Intellectual Disability	N=10, not analysed	N=16, Fisher's exact = 10.266, p=.017, df=4*
Specific Learning Disability – WR	N=10, not analysed	N=11, not analysed
Specific Learning Disability – Sp	N=14, not analysed	N=16, Fisher's exact=14.348, p=.001, df=4. *
Eneuresis or Encopresis	N=2, not analysed	N=7, not analysed

8.3. Referral-Defined and Team-Defined Problems

An equivalent analysis of referral problems identified that in the cases where the number of cases was sufficient to conduct statistical analyses, there were no significant changes in referral problems over the time periods studied (see Table 2.15). Again, this suggests that there has been no clear change in the problems detailed in the referrals received by either service within CAFT over the three time periods.

Table 2.15. A summary of analyses conducted on referral problems over by time period and clinic.

Referral Problem	CPS	A&FS
Attachment Difficulties	N=3, not analysed	N=31, $\chi^2=3.392$, p=.195, df=2
Conduct or ODD Symptoms	N=89, $\chi^2=5.567$, p=.237, df=4	N=89, $\chi^2=1.815$, p=.775, df=4
Callous/Unemotional Traits	N=3, not analysed	N=6, not analysed
ADHD Symptoms	N=34, $\chi^2=.492$, p=.809, df=2	N=30, $\chi^2=4.989$, p=.087, df=2
ASD Symptoms	N=3, not analysed	N=14, not analysed
Anxiety Symptoms	N=5, not analysed	N=17, Fisher's exact = .166, p=1, df=2
Mood Disorder Symptoms	N=10, not analysed	N=6, not analysed
Suspected Intellectual Disability	N=1, not analysed	N=6, not analysed
Eneuresis or Encopresis	N=4, not analysed	N=11, not analysed

The analysis conducted for presenting problems showed similarly low numbers of disorders other than CD/ODD and ADHD, and again where the number of cases was sufficient to conduct statistical analyses, there were no significant changes in presenting problems over the time periods studied (see Table 2.16). This suggests some stability in the presenting problems of children seen by CAFT over the time periods studied.

Table 2.16. A summary of analyses conducted on presenting problems over three time periods per clinic.

Presenting Problem	CPS	A&FS
Attachment Difficulties	N=0, not analysed	N=19, $\chi^2=4.367$, p=.107, df=2
Conduct or ODD Symptoms	N=89, $\chi^2=5.505$, p=.246, df=4	N=100, Fisher's exact = 1.412, p=.855, df= 4
ADHD Symptoms	N=42, $\chi^2=3.081$, p=.228, df=2	N=50, $\chi^2=3.900$, p=.143, df=2
ASD Symptoms	N=14, not analysed	N=19 Fisher's exact = 4.967, p=.254, df=2
Anxiety Symptoms	N=7, not analysed	N=16, Fisher's exact = 1.026, p=.594, df=2
Mood Disorder Symptoms	N=11, not analysed	N=11, not analysed
Suspected Intellectual Disability	N=1, not analysed	N=3, not analysed
Eneuresis or Encopresis	N=8, not analysed	N=16, $\chi^2=1.956$, p=.442, df=2

8.4. Parent- and Teacher-Rated SDQ Scores

8.4.1. Parent-Rated Strengths and Difficulties

The parent-rated SDQ total and subscale scores were entered as dependent variables within a ANOVA on the factors Time Period (3 levels) and Service (2 levels). Neither the main effect of Time Period ($F(12,278)=1.235$, $p=.258$, 70% power), Service ($F(6,138)=1.202$, $p=.309$, 46% power), or their interaction ($F(12,278)=.924$, $p=.523$, 54% power), were significant, and so no further analyses were conducted. The descriptive statistics for this analysis are presented in Table 2.17.

8.4.2. Teacher-Rated Strengths and Difficulties

The teacher-rated SDQ total and relevant subscale scores were entered as dependent variables within a MANOVA on the factors Time Period (3 levels) and Service (2 levels). Neither the main effect of Time Period ($F(6,150)=.664$, $p=.679$, 26% power), Service ($F(3,74)=.796$, $p=.500$, 21% power), or their interaction ($F(6,150)=.623$, $p=.712$, 24% power), were significant and so no further analyses were conducted. The descriptive statistics for this analysis are presented in Table 2.17.

Table 2.17. Descriptive statistics for SDQs by Service and Time Period.

Service	Time Period	Statistic	SDQ-T			SDQ-P		
			Total	Behaviour	Hyperactivity	Total	Behaviour	Hyperactivity
CPS	2 years pre	N	7	7	7	23	25	25
		Min	12	2	5	9	3	2
		Max	26	8	10	30	10	10
		Mean	17.00	4.71	7.86	21.96	6.12	7.96
		SD	4.546	2.215	1.773	6.449	2.242	2.091
		SEM	1.718	.837	.670	1.345	.448	.418
	0-2 years post	N	12	12	12	25	24	24
		Min	8	1	1	11	3	4
		Max	30	9	10	36	10	10
		Mean	17.58	4.08	6.83	23.36	6.50	7.83
		SD	7.937	2.193	3.186	7.199	2.246	2.259
		SEM	2.291	.633	.920	1.440	.458	.461
	2-4 years post	N	21	21	21	25	25	25
		Min	2	0	1	10	0	3
		Max	32	9	10	37	10	19
		Mean	19.33	5.48	7.19	22.00	6.12	8.12
		SD	9.140	3.530	2.960	7.621	2.571	3.270
		SEM	1.994	.770	.646	1.524	.514	.654
A&FS	2 years pre	N	0	0	0	1	0	1
		Min	35	10	10	32	9	10
		Max	17.31	4.00	7.00	18.07	4.90	6.31
		Mean	8.631	3.062	3.882	8.229	2.440	3.230
		SD	16	17	16	29	30	29
		SEM	2.158	.743	.970	1.528	.446	.600
	0-2 years post	N	13	13	13	33	33	33
		Min	11	0	3	8	0	1
		Max	30	8	10	36	10	10
		Mean	19.15	4.00	7.92	22.52	5.97	8.00
		SD	5.728	2.236	2.929	6.974	2.481	2.512
		SEM	1.589	.620	.812	1.214	.432	.437
	2-4 years post	N	13	13	13	15	16	16
		Min	9	0	2	4	0	4
		Max	29	9	10	35	10	10
		Mean	17.69	4.31	8.15	18.67	5.37	7.56
		SD	6.447	2.955	2.672	8.338	3.442	2.097
		SEM	1.788	.820	.741	2.153	.861	.524

8.5. Psychometric Assessment Results

The variables Verbal IQ, Performance IQ, Full-Scale IQ, Spelling Standard Score and Word Reading Standard Score were entered into an ANOVA on the factors Time Period (3 levels) and Service (2 levels). Neither the main effect of Time Period ($F(10,184)=1.269$, $p=.251$, 64% power), Service ($F(5,91)=1.629$, $p=.160$, 54% power), or their interaction ($F(10,184)=1.268$, $p=.251$, 64% power) Automatically computed univariate tests did, however, show significant effects on measures of performance IQ ($F(1,95)=5.190$, $p=.025$, partial eta squared=.052, 61% power) and full-scale IQ ($F(1,95)=5.969$, $p=.016$, partial eta squared=.059, 68% power). These differences are not of particular relevance, as they reflect the increased number of cases with learning disabilities within the A&FS during the second time point relative to the other two time points. See Table 2.18. for the relevant descriptive statistics. The only significant differences related to a decrease in IQ scores during the middle time period in the A&FS, reflected the increased incidence of Learning Disabilities in this cell.

Table 2.18. Descriptive Statistics of Psychometric Assessment Scores by Service and Time Period.

Service	Time Period	Statistic	VIQ/VCI	PIQ/PRI	FSIQ	Word Reading	Spelling
CPS	2 years pre	N	28	27	21	16	14
		Min	68	52	63	72	76
		Max	127	144	124	152	122
		Mean	101.04	94.96	99.38	107.69	98.93
		SD	16.756	21.063	18.329	20.584	13.373
		SEM	3.167	4.054	4.000	5.146	3.574
	0-2 years post	N	27	27	21	21	18
		Min	55	57	59	40	40
		Max	134	135	139	139	139
		Mean	90.70	88.74	90.05	86.76	82.28
		SD	19.574	19.461	21.681	23.274	20.332
		SEM	3.767	3.745	4.731	5.079	4.792
	2-4 years post	N	26	26	19	21	20
		Min	73	70	70	55	52
		Max	130	126	132	124	117
		Mean	94.88	94.46	96.11	89.52	85.20
		SD	15.381	15.166	17.304	20.366	15.936
		SEM	3.017	2.974	3.970	4.444	3.563
A&FS	2 years pre	N	35	34	27	28	24
		Min	59	60	0	50	58
		Max	130	131	115	112	113
		Mean	93.54	86.47	85.63	83.21	82.96
		SD	16.630	14.888	23.107	16.222	16.212
		SEM	2.811	2.553	4.447	3.066	3.309
	0-2 years post	N	45	45	39	31	30
		Min	57	64	57	50	58
		Max	191	131	123	114	114
		Mean	93.02	86.78	87.31	85.16	83.40
		SD	21.765	14.121	15.192	15.614	12.571
		SEM	3.244	2.105	2.433	2.804	2.295
	2-4 years post	N	23	21	18	21	19
		Min	70	63	72	58	47
		Max	123	133	121	123	123
		Mean	94.09	95.33	94.00	92.71	88.89
		SD	13.259	19.363	17.597	18.626	18.935
		SEM	2.765	4.225	4.148	4.064	4.344

9. Appendix 3: Supplementary Information on Primary Analyses

Table 2.19. Descriptive statistics of the complexity index by Service and Time Period.

Note that complexity index scores have a range of 0-5.

Service	Time Period	N	Mean	SD
CPS	2 year pre	31	1.7097	.90161
	0-2 years post	33	2.3636	1.14067
	2-4 years post	38	2.4211	1.03013
	Total	102	2.1863	1.06919
A&FS	Before	40	1.5000	1.17670
	0-2 years post	50	1.6800	.99877
	2-4 years post	35	1.5714	1.03713
	Total	125	1.5920	1.06329

Table 2.20. Number of treatments planned for each service in each time period.

Service	Time Period	N	Mean	SD
CPS	2 year pre	33	1.94	1.22
	0-2 years post	33	2.45	1.25
	2-4 years post	39	2.38	1.46
	Total	105	2.26	1.33
A&FS	Before	43	2.14	.97
	0-2 years post	50	2.52	1.37
	2-4 years post	34	1.97	1.38
	Total	127	2.24	1.26